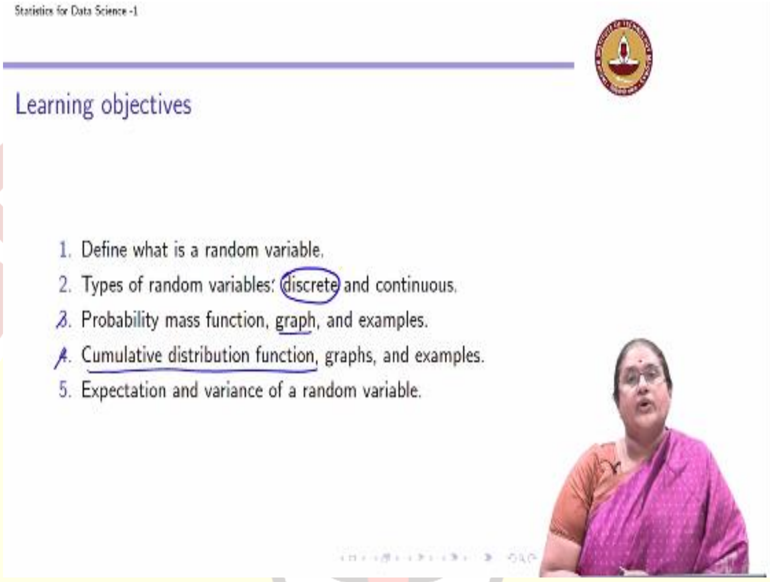
IIT Madras

ONLINE DEGREE

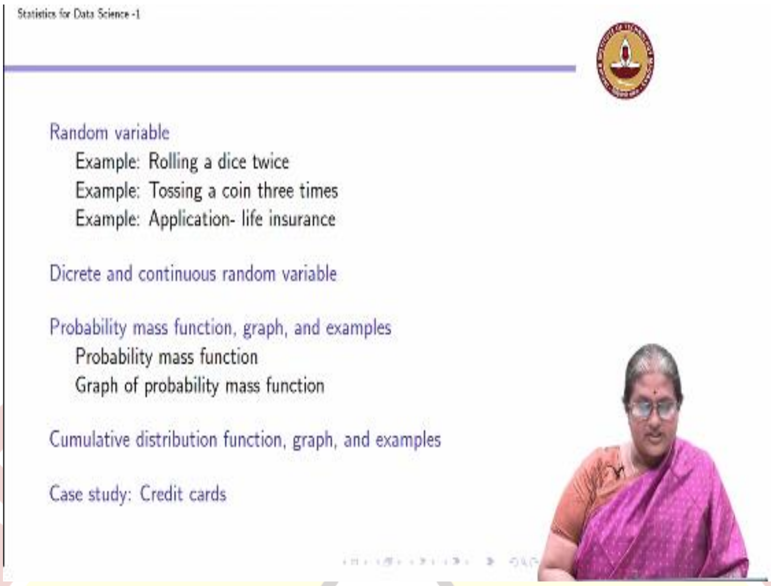
Statistics for Data Science-1
Professor. Usha Mohan
Department of Management Studies
Indian Institute of Technology, Madras
Lecture No. 8.5
Discrete Random Variables - Graph of Probability Mass Function

(Refer Slide Time: 00:14)



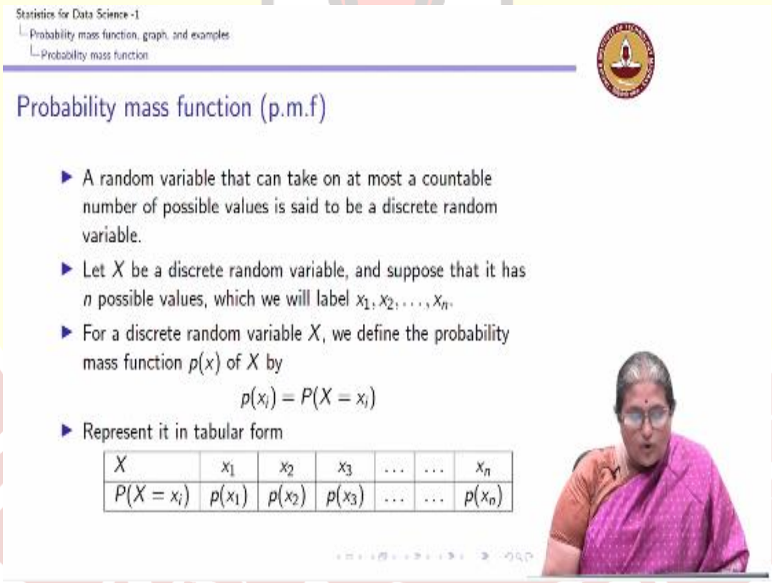
So, so far what we have seen is we defined what is a random variable and then we discussed about discrete and continuous random variables and we defined what was a probability mass function. So, today we are going to continue to understand about what is a probability mass function, we will see what do we mean by a graph of a probability mass function and then we will introduce what is a cumulative distribution function all of this when the variable is a discrete variable that is the learning objective.

(Refer Slide Time: 00:50)



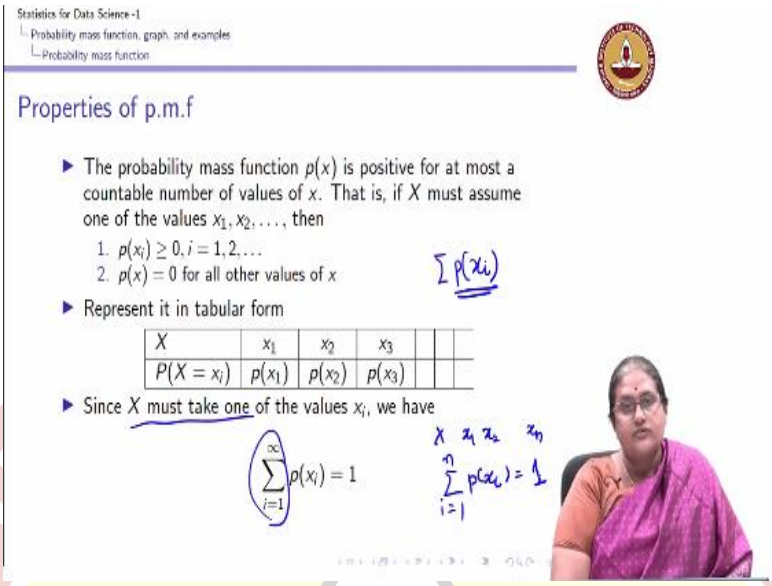
So, let us go and review what we have learned about a probability mass function so far.

(Refer Slide Time: 00:57)



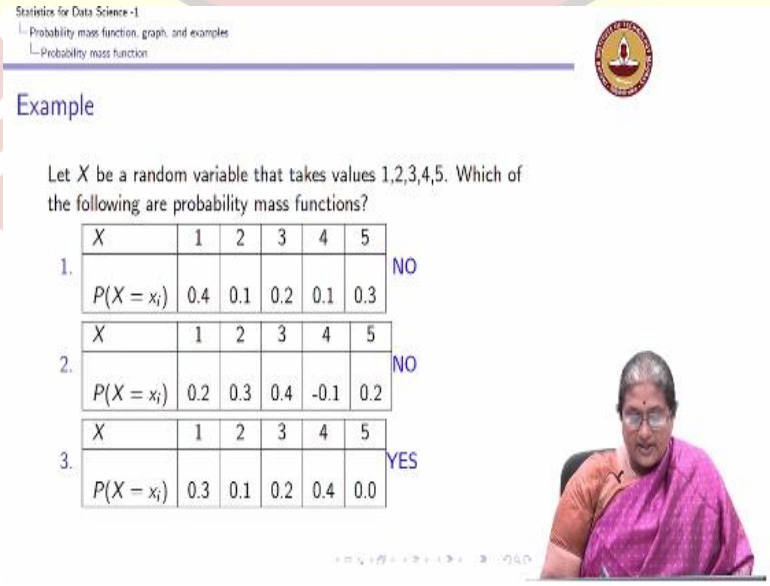
We saw that a probability mass function of a discrete random variable is a random variable which can take at most countable number of possible values. Now if each of these random variables which we are labeling them as x_1, x_2 that is X is taking these values the probability mass function is a probability with which the random variable X takes a particular value x_i.

(Refer Slide Time: 01:35)



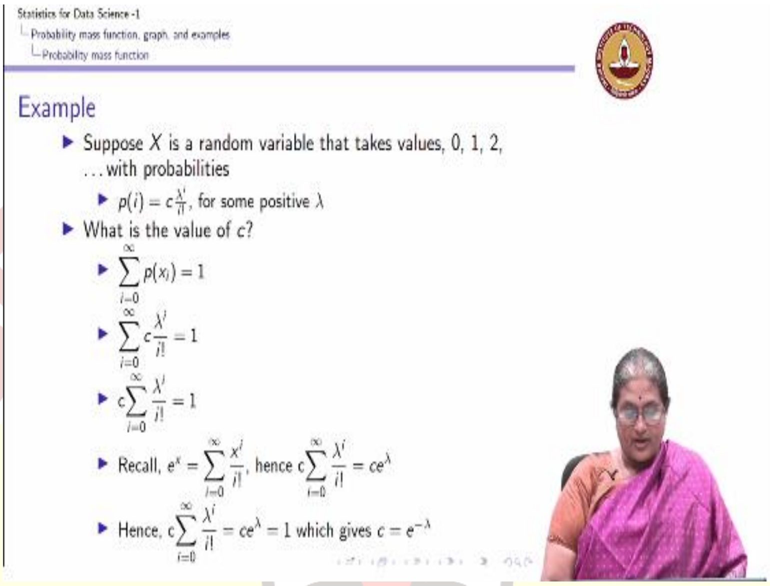
Now since this random variables takes these values, then we know the properties of a probability mass function is that probability since we are talking about a probability, $P(X_i)$ in other words probability of X equal to xi is always non-negative and the next important property is since X must take one of the values summation over all possible values of x. If X takes the values x_1, x_2 up to $\sum_{i=0}^{\infty} P(x_i) = 1$if X takes the value x_1, x_2, x_n finite number of values, then it would be $\sum_{i=0}^{n} P(x_i) = 1$. So, these are the two important and critical properties of a probability mass function.

(Refer Slide Time: 02:32)



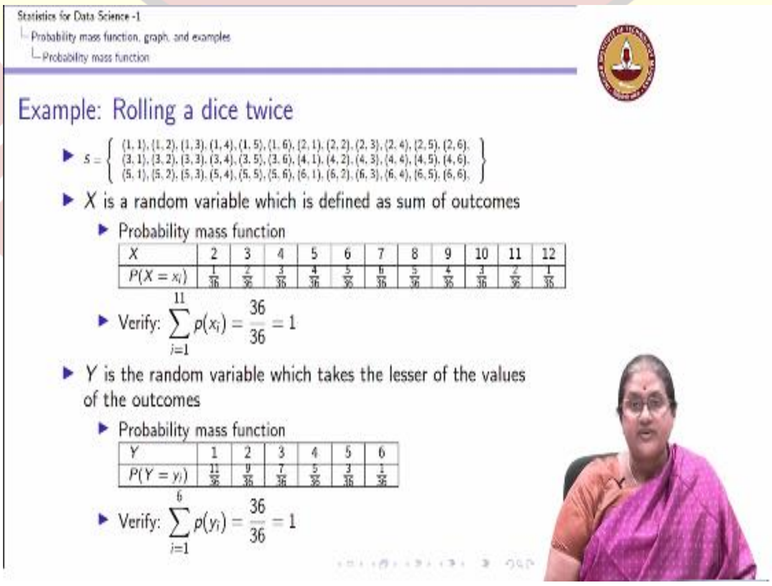
What should we be able to answer? So, we should be able to answer given a probability mass function we should be able to verify whether something is a probability mass function. This we achieve by checking out all the properties.

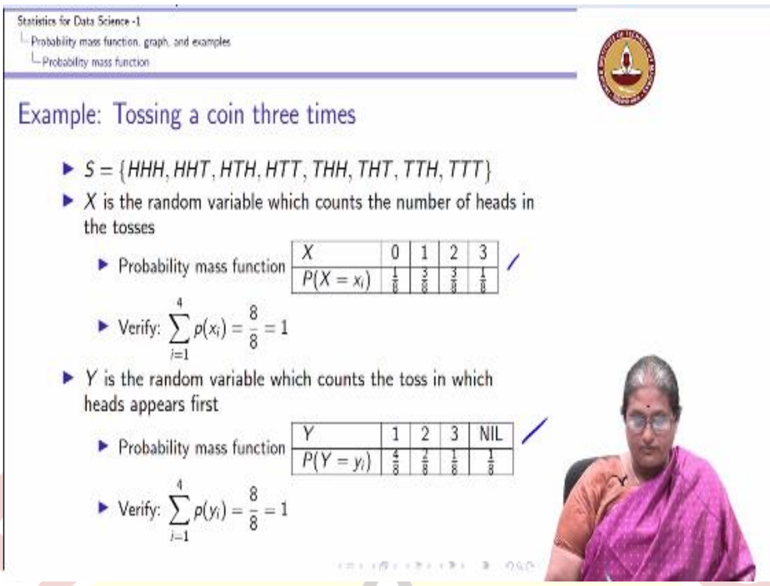
(Refer Slide Time: 02:49)



And then afterwards suppose I have given a probability mass function where I have to find out a constant then we also saw that by equating this to one we can find out what is the value of the constant for which this would define a probability mass function.

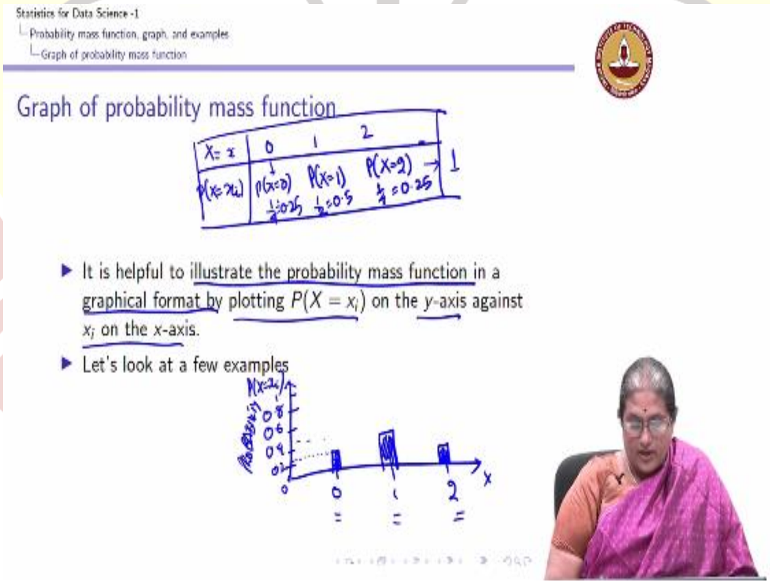
(Refer Slide Time: 03:14)





So, this is what we have done in our earlier case. We went back to the examples that is rolling a dice twice to see what is the tabular form of the probability mass function for all the random variables we have defined. Both in rolling a dice twice and tossing a coin thrice. So, this is what we have seen we have seen these were the tabular form of actually illustrating or presenting my probability mass function.

(Refer Slide Time: 03:44)



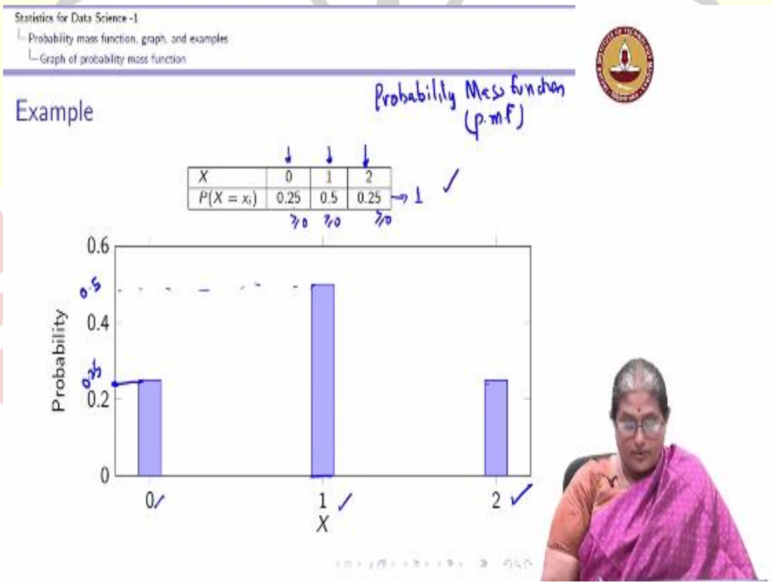
Now we go to the next thing of how can I graph the probability mass function? Now why should I graph the probability mass function? It is always helpful to illustrate the probability mass function in a graphical format. What do I mean by a graphical format? Recall, a probability mass function has the values of X for example if X takes the value 0, 1, 2 then what I have in a tabular form is this would give me what is the probability of P(X=0).

This would give me what is the P(X=1) and this will give me P(X=2) I know the sum of all these probabilities would be 1 and since they are probabilities all of them are greater or equal to 0. So, the question is can I represent or illustrate this information in the form of a graph. So, what do we mean by that? I can plot the probabilities on the y axis against x_i on the x axis.

So, a simple way to discuss this is what are the values X takes here? X takes the value 0, 1 and 2. Remember that these are just values that is taken if they are ordinal then the order has to be maintained on the y axis I am going to have the probability of X taking a particular value. I just write it as probability. So, this is what is my probability mass function. So, suppose probability X equal to 0 is a one fourth, X equal to 1 as a half, I know that the y axis takes values say 0.2, 0.4, 0.6, 0.8 and 1 because it is a probability it would start with a 0 and end with a 1.

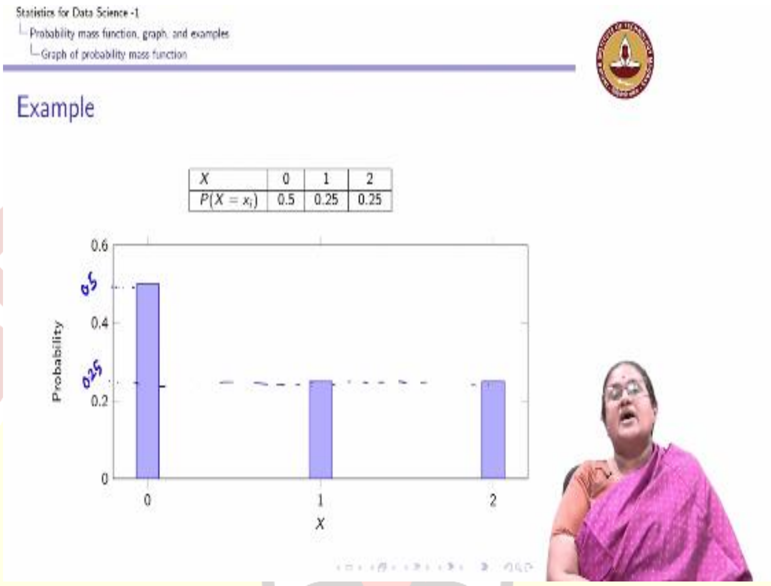
So, P(X=0)=0.25 which is 1/4, 1/2 is a 0.5 again 1/4 which is again a 0.25. I can construct the probability mass function I plot probability so this is a 0.25 so I can go here I construct a bar which is 0.25. Similarly, this is 0.5 P(X=1)=0.5 again P(X=2) = 0.25.

(Refer Slide Time: 06:31)



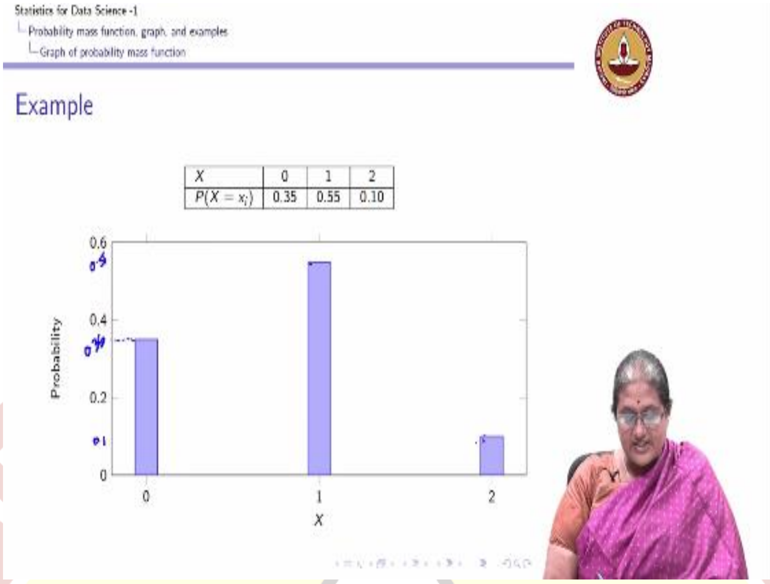
So, let us look at this example so I have this is a tabular form of my probability mass function. I have X equal to 0 is 0.25, 0.5 and 0.25. So, X takes the value 0, 1 and 2. This would be 0.25 so this corresponds to 0.25 this height corresponds to 0.5 this is again corresponding to 0.25. So, this is what we refer to as the probability mass function. It is abbreviated as pmf and the reason it is a probability mass function as we can imagine this as a discrete points that x takes. I have a weight associated and this is what is at every discrete point 0.25, 0.5 and 0.25. Again I can verify that this adds up to 1 so it is a probability mass function all of them are non negative so I do not have a problem.

(Refer Slide Time: 07:56)



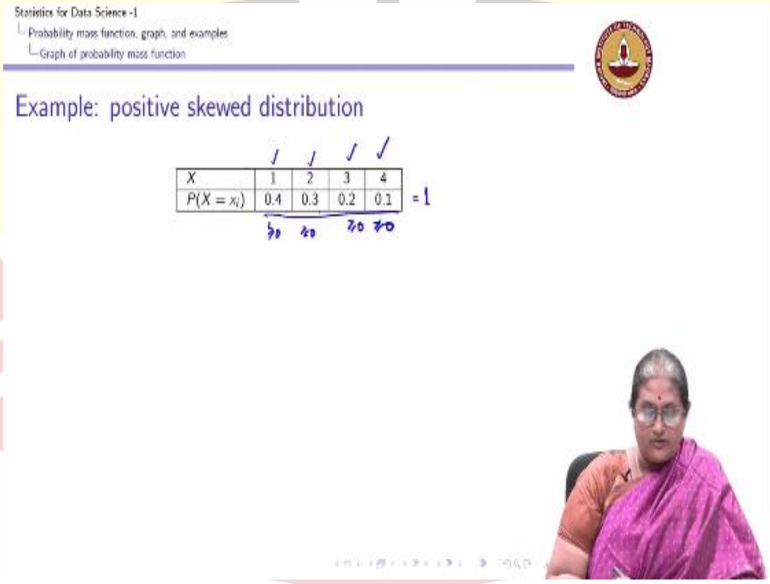
Now let us look into another example. Here again I verify this is 0.5 plus 0.25 plus 0.25 it adds up to 1 all of them are greater or equal to 0 so it is a probability mass function. So, in this case I can see that probability X equal to 0 takes the value 0.5, X equal to 1 takes the value 0.25 so is X equal to 2 so this is my probability mass function. What you can notice here is in the earlier example the distribution looks symmetric whereas here I do not see a symmetricity in my distribution.

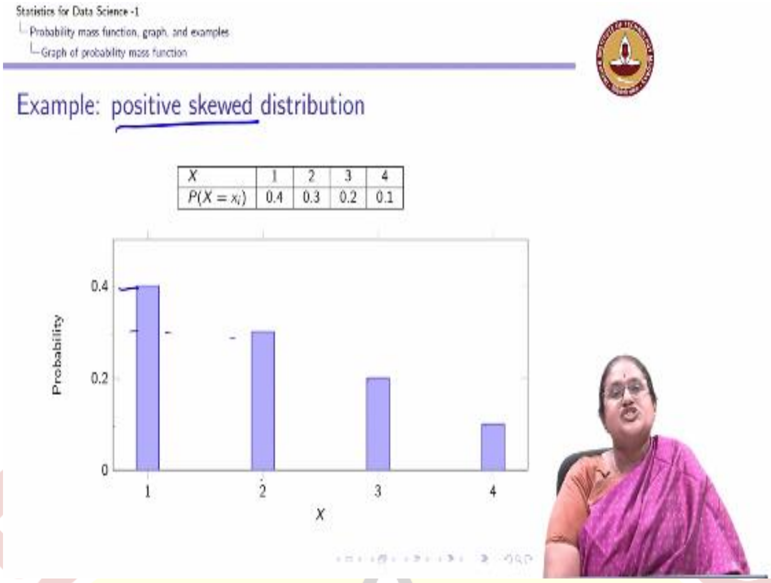
(Refer Slide Time: 08:41)



I could also have a case where I do not have any pattern I have here X equal to 0 is 0.35, X equal to 1 is 0.55 and X equal to 2 is a 0.1 and that is demonstrated or illustrated by this graph. So, why is a probability mass function important or useful?

(Refer Slide Time: 09:15)

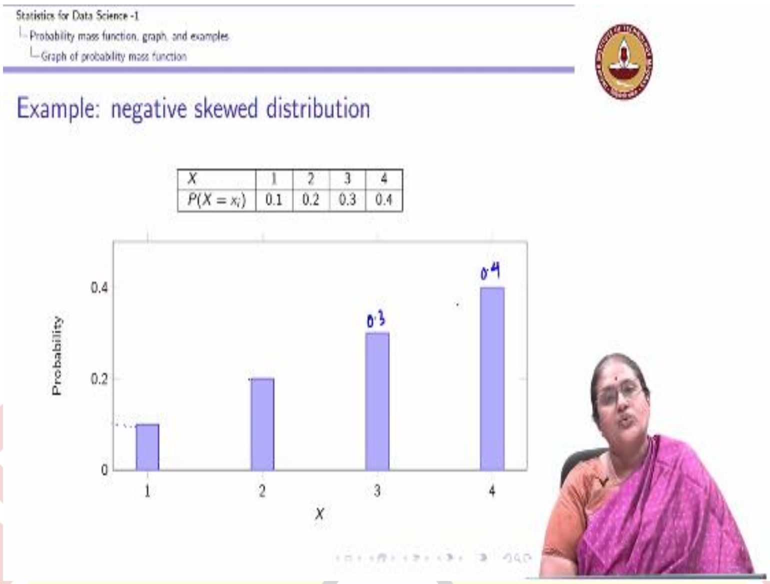




Sometimes we can see that this gives us the shape of the distribution of the random variable. What do we mean by this? For example, if I have this case where X is again is a discrete random variable which takes the value 1, 2, 3, 4 with these probabilities again 0.4 + 0.3 is 0.7, 0.7 + 0.2 is 0.9, 09 + 0.1 all of them add up to 1 and all of them are non negative. So, I have a probability mass function.

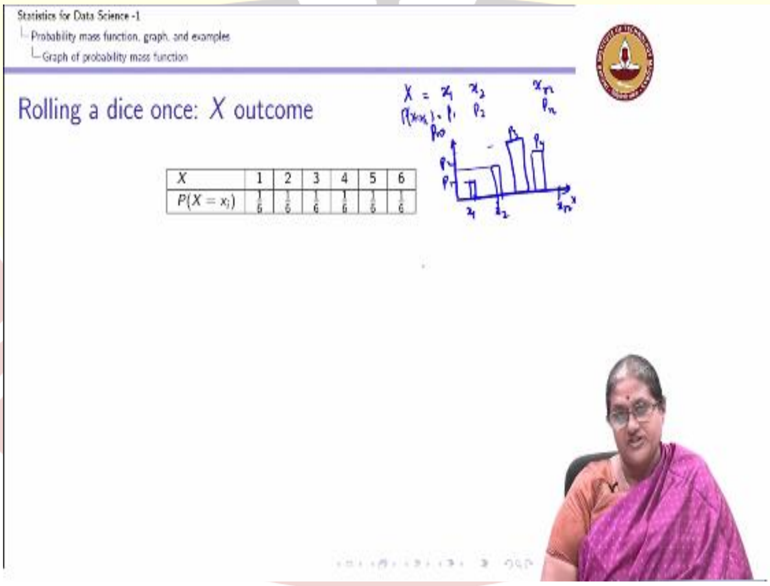
Now if I plot this I can see that this is a 0.4, X equal to 2 is a 0.3, X equal to 3 is a 0.2 and X equal to 4 is a 0.1. What you can see is this distribution exhibits a skewness and these are referred to as positive skewed distributions. So, in advanced courses in addition to the center and variability the shape of a distribution is very important. So, you can see that this is a skewed distribution.

(Refer Slide Time: 10:36)



Now if you look at the other example for example here my 1 probability X equal to 1 is a 0.1, X equal to 2 is a 0.2, X equal to 3 is a 0.3 and X equal to 4 is a 0.4 you have what we refer to as a negative skewed distribution.

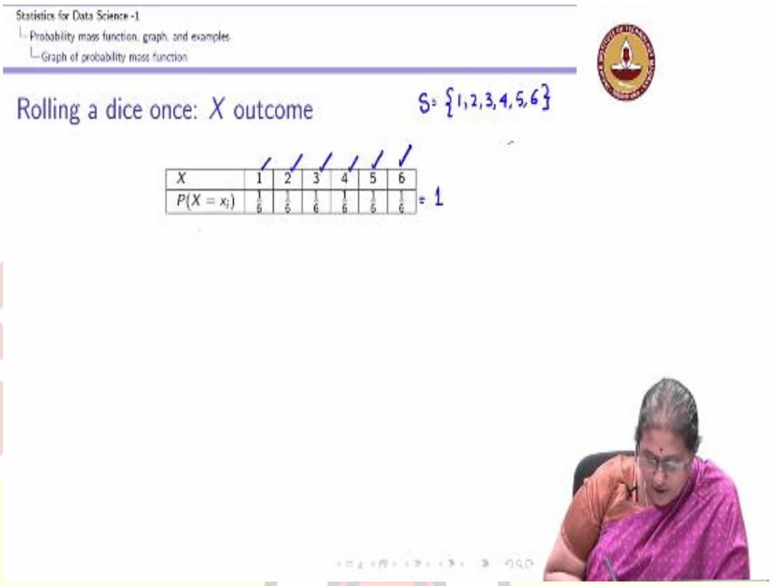
(Refer Slide Time: 10:59)



So, in summary you can see that whenever I have a distribution of x taking values x_1, x_2, x_n with probability of X equal to x_i this p_1, p_2, p_n. For now I am assuming X takes finite number of values, then the graph which plots x_1, x_2, and x_n on the x axis and the probabilities that is this would be corresponding to p_1 this could be p_2 this could be p_3 and this could be my p_4 and so forth this is referred to as a graph of a probability mass function.

So that is what we have seen so far. So, now let us go back to the examples that we have considered and look at how the PMF can be illustrated using a graph.

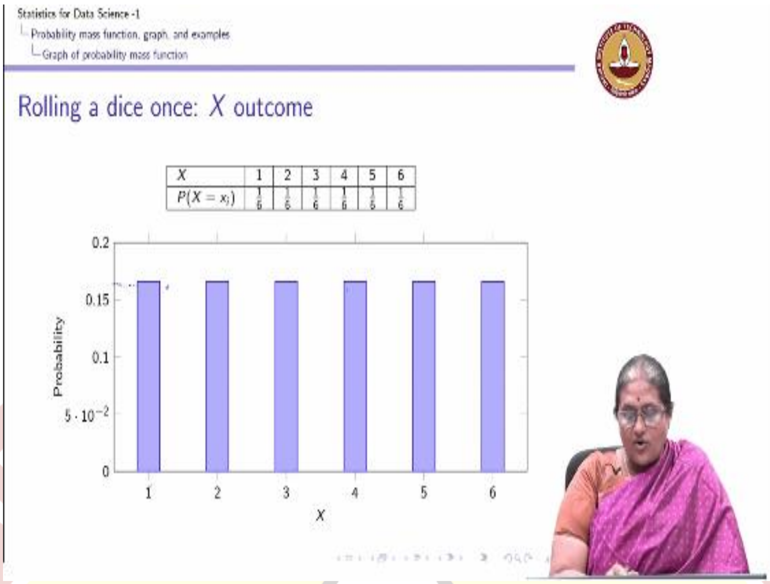
(Refer Slide Time: 12:08)



So, let us start with rolling a dice 1 I know the sample space for this is going to be the following. I am rolling a 6 sided dice it is a fair dice so I am rolling it only once. I know that the outcomes could be any one of the 6 outcomes it is a fair dice. So, the random variable takes the values 1, 2, 3, 4, 5 and 6 with the probabilities because it is a fair dice the chance of getting a 1 is the same as the chance of getting a 2 and so forth.

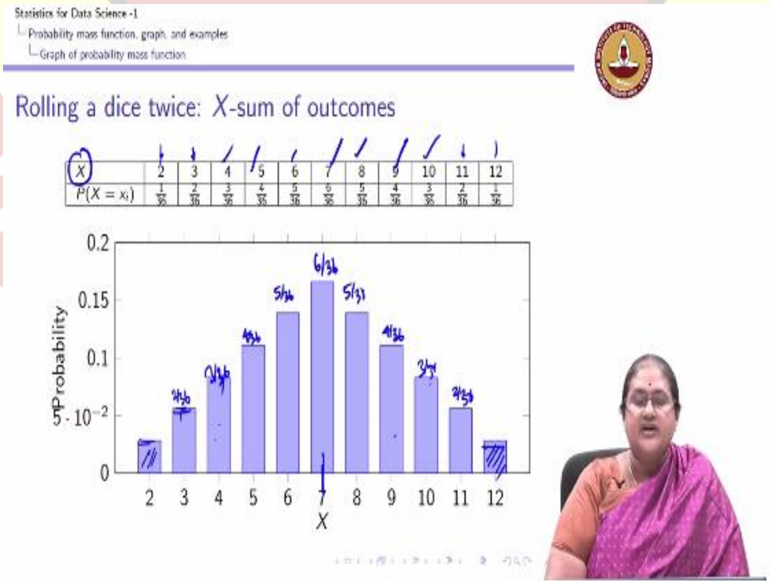
So, I have listed the probability of X taking the values 1 to 6 is the same which is 1 by 6, 1 by 6, 1 by 6 so forth. We can again verify that the sum is equal to 1 all of them are greater than or equal to 0 all the probabilities are now negative. Hence it defines a probability mass function.

(Refer Slide Time: 13:17)



Now if I plot this probability mass function I can see that it is a constant you can see that for every value X takes the height of the bars are the same which is around 1.16. So, this is one shape so the graph of a probability mass function illustrates that there is some degree of uniformity in my distribution. So, immediately you can see that this looks like a uniform distribution. We will talk about distributions later, but what you can see from the graph is immediately you can see all the bars are of the same height.

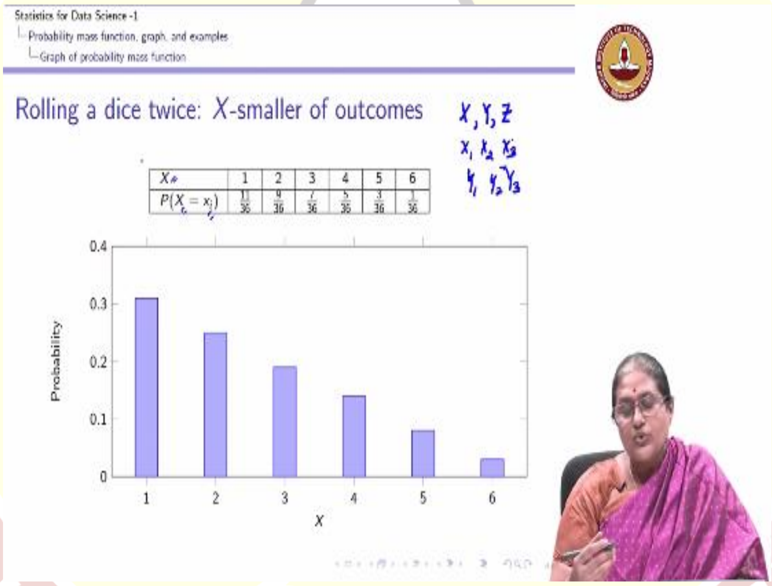
(Refer Slide Time: 14:06)



Now let us go back to rolling a dice twice. Again here what is a random variable? Random variable is the sum of outcomes so we have already seen that this is the tabular form X takes values 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12 with the respective probabilities as shown in the table. Now if I come and look at the probability mass function you can see that X takes the value 2 with a very, very low probability. X takes the value 3, but X takes the value 2 and 12 with the same probability as shown here by the bars.

Similarly, X X takes the value 3 and 11 with the same probability which is 2 by 36 it takes the value 4 and 10 with the value 3 by 36 it takes the values 5 and 9 with the value 4 by 36 probability. It takes the value 6 and 8 with the probability is 5 by 36 and it takes the value 7 with the probability 6 by 36. So, again you can see that the distribution of this random variable is symmetric about a particular point. You can see that there is a element of symmetricity or you can see a symmetric behavior in the distribution.

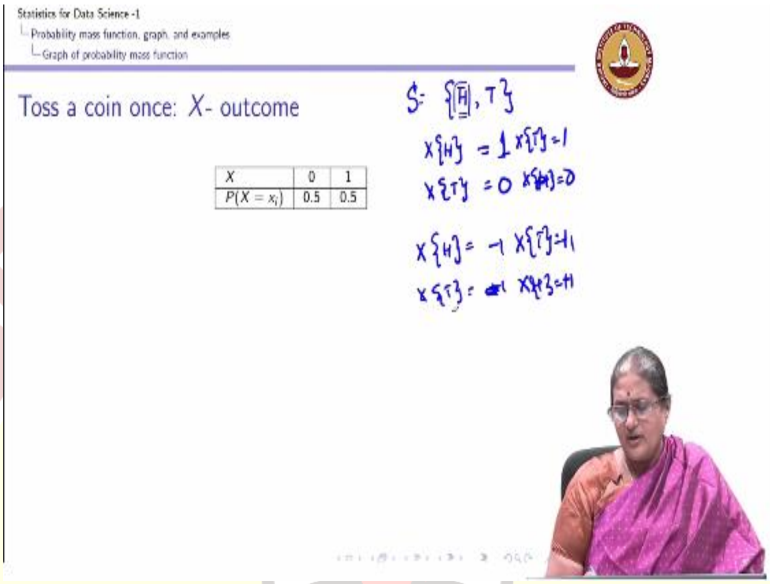
(Refer Slide Time: 15:50)



Now let us look at smaller of the outcomes again we know that X takes values 1, 2, 3, 4, 5, 6 the sum adds up to 1 once I plot this I know X takes these values you can see that it is a skewed distribution. Again recall, both these random variables were defined on the same random experiment and sample space. It depends on what is it you are actually mapping each of the outcomes to and you can see that the distribution in one case was a symmetric distribution in the other case it was a skewed distribution.

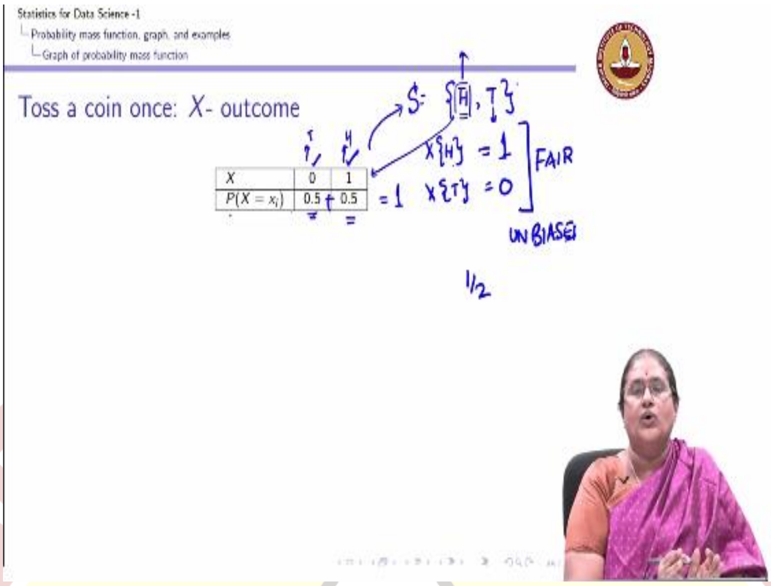
I just want to make a small point here when we discuss this random variable I had defined it as a y and I said probability Y equal to yi here I am referring it to as X and I am looking at X equal to xi. It really does not make any difference whether you are referring it to X or Y only thing you need to understand that what represents the random variables and what represents the value they are of the random variable. It could be X, Y, Z. So, typically we also saw that random variables are typically expressed with upper case alphabets X, Y, Z or X_1, X_2, X_3 or Y1, Y2, Y3 these are typically what are used to represent random variables.

(Refer Slide Time: 17:37)



Now let us look at the tossing a coin once. Now again what is a sample space for my experiment? When I am tossing a coin once I can get a head or a tail. Now again if I want to just note the outcomes I would associate a value to these outcomes. There are two ways I can do this mapping, either I can map head to 1, tail to 0 or I can map a tail to 1 or head to 0. So, you can see that a random variable just maps the value of an outcome to a number. Can we map head to a minus 1 and tail to a plus 1 or tail to a minus 1 and head to a plus 1 we can do it.

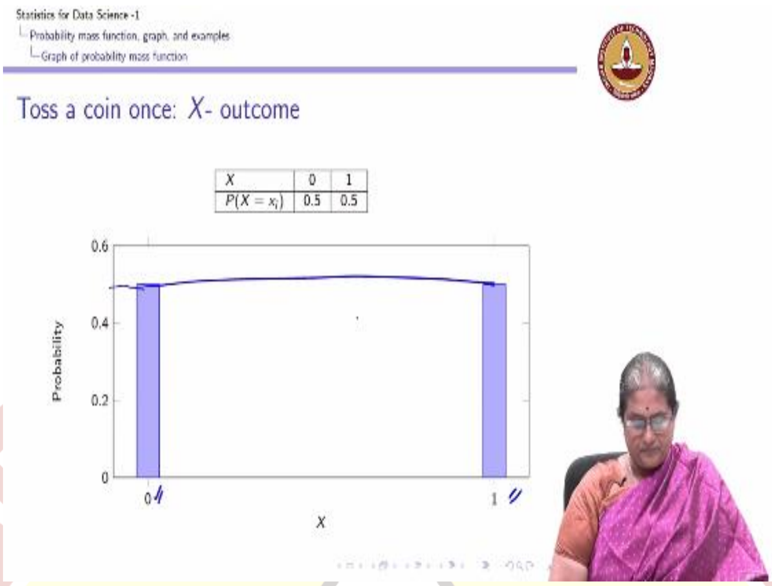
(Refer Slide Time: 18:41)



But for convention sake let us always for just our understanding now we will map head to 1 and tail to 0. Whatever I am mapping I have to define it very clearly. So, if head takes the value 1, I know the random variable again takes the value 0 and 1. Again, I assume that my coin is a fair coin or an unbiased coin. If I have a fair coin or an unbiased coin, the probability of getting a head is the same as probability of getting a tail which we know is half.

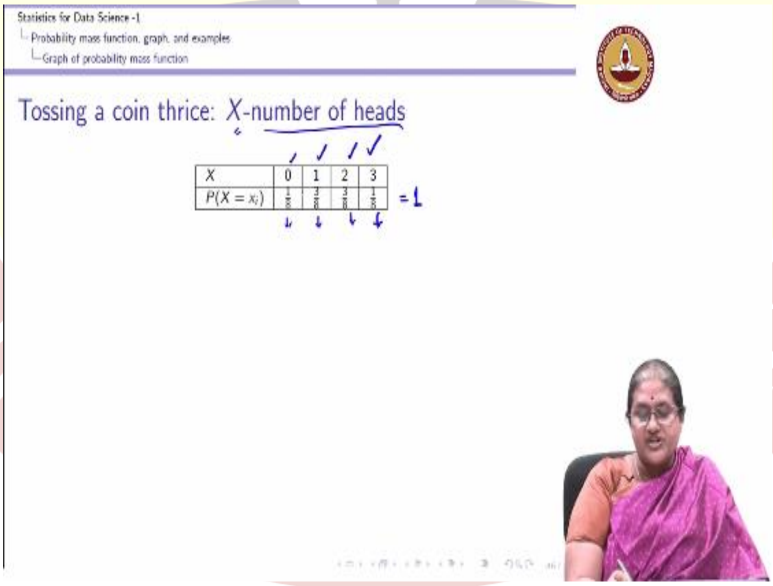
This is something which I know the probability of getting a head is the same as probability of getting a tail which is half so I am just putting here probability X equal to xi is 0.5, 0.5 both of them I add that becomes 1. This for now is a tail and this is a head. Whenever we are doing this mapping we need to understand what was our original experiment and what has been the mapping or how we have associated every outcome to a random variable that is something which we need to know.

(Refer Slide Time: 20:09)



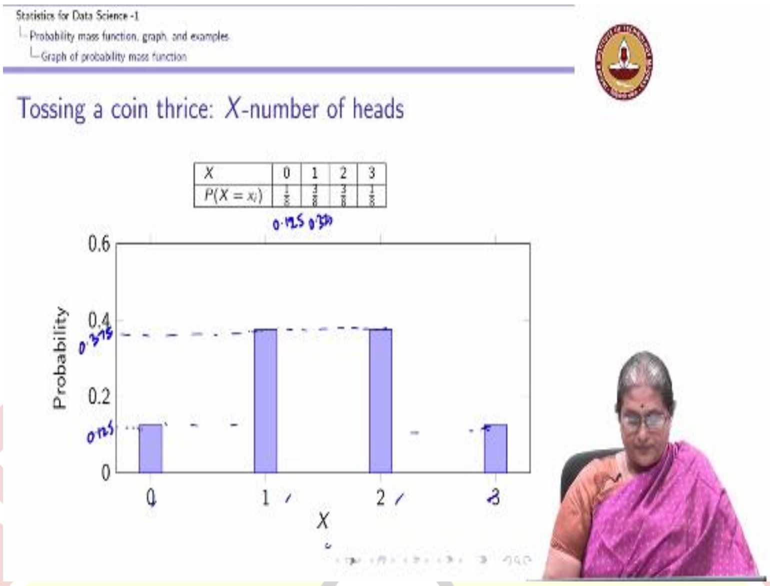
Now given this x takes only two values 0 and 1 with the same probability again we can see that this probability since both of them have the same probability there is some sort of a constant or a uniformity which is illustrated here.

(Refer Slide Time: 20:31)



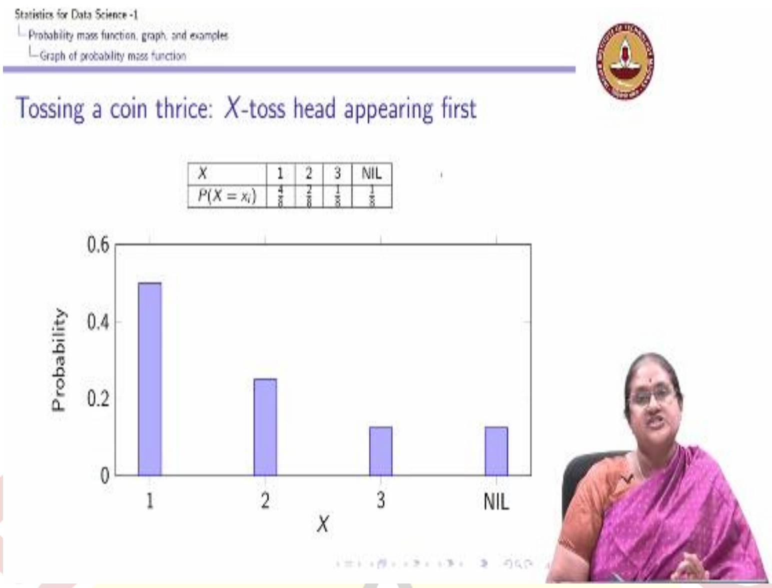
Let us go to the example of tossing a coin 3 times again my random variable was counting the number of heads we knew that the value this random variable takes is 0, 1, 2 or 3 with the probability 1 by 8, 3 by 8, 3 by 8 and 1 by 8. We again verified this is a probability mass function because the sum of the probability is equal to 1.

(Refer Slide Time: 20:59)



Now again if I plot this I can see that X takes the value 0, 1, 2, 3 with probabilities 1 by 8 is 0.125. This is 0.375 so 0.125 is here, 0.375 is here again 2 takes 0.375, X equal to 3 is 0.125 this is 0.375 again you see that this distribution is a symmetric distribution around the values X takes. You can see from a graph there is some amount or it indicates symmetric distribution.

(Refer Slide Time: 21:45)



Now let us look at the next example of tossing a coin thrice where I am seeing the toss where head appears first. Again we have seen already that head can take the values, X can take or 1, 2, 3 in other words a head can appear first in the first toss or the second toss or the third toss we defined nil. Technically, this nil is not a real number we should map this nil to a real number we also mentioned about this earlier, but I am just going to retain the nil now.

And then you can see that the distribution of this random variable again exhibit some sort of skewness as we have already discussed earlier what is a skewed distribution. So, you can see that going back to the examples that we have already learned whenever we are talking about the probability mass function we can describe the distribution through the shape of a distribution. So, the next thing which we are interested in knowing is what we mean by a cumulative distribution function.