Mine Automation and Data Analytics

Prof. Radhakanta Koner

Department of Mining Engineering

IIT (ISM) Dhanbad

Week - 08

Lecture 37: Discrete Random Variable Part I

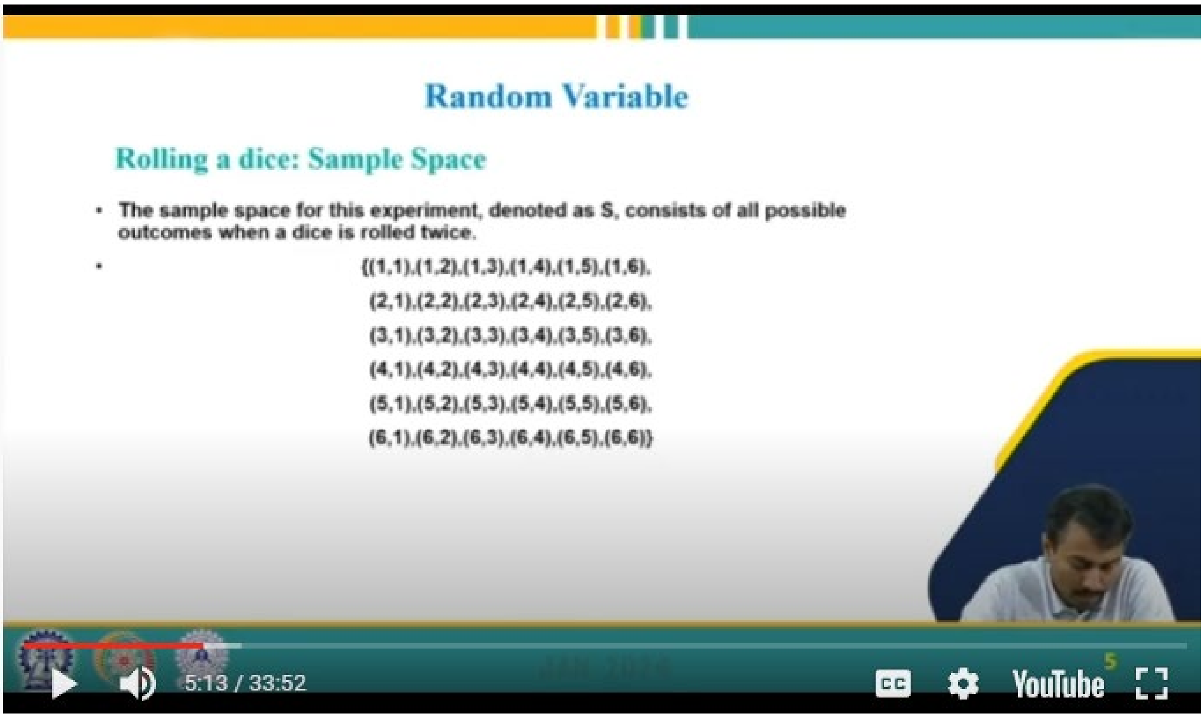
Welcome back to my course, mine automation and data analytics. Today, we will discuss a discrete random variable that is a part of inferential statistics. So, under this, we will discuss the definition of random variables, and we will give some examples and then types of random variables that are discrete and random. Then, we will discuss the probability mass function with graphs and examples, cumulative distribution functions, an important concept we will discuss, expectation and its properties, and variance. So, what is a random variable? So, a random variable is an outcome of an experiment. It quantifies the result of the experiment.

For example, we flip the coin several times in an experiment. So, we continuously flip the coin, for example, ten or 15 times. So, we may note the number of times two heads appeared. So, two heads in the countable number, for example, two heads appeared three times out of 10.

So, the probability of that particular is three by 10. So, we are looking at how often two heads appear continuously for that experiment. So, for example, another is rolling the dice that two times you are rolling the dice, one then another one, for example, two dice is you are moving. So, what is the chance that the sum of the rolling dice is 7? So, you have to select from the total space the possibility of rolling two dice, the possible set and space, and how many times the summation of these two outcomes is 7.

So, 7 is a round number, and the result is the summation. So, at different times, we could be more precise about what is happening, but some outcomes may be needed for some real examples. So, these random variables are a handy example. So, here we have enumerated that. So, out of that, how many times are there seven coming?

So, a combination of 1, 6, 2, 5, 3, 4, 5, 2, 6, 1 is a total of 7, which is the total of rolling dice two times two rolling dice. So, these quantities of interest or, more formally, these real-valued functions defined on the sample space are called random variables. Since the experiment's outcome determines the value of a random variable, we can assign probability to the possible values of the random experiment. So, the sample space of the experiment is denoted as S and consists of all possible outcomes when the dice are rolled twice. So, 1, 1, 1, 2, 1, 3, 1, 4, 1, 5, 1, 6 till 6, 6.



So, 1, 2, 3, 4, 5, 6, and 6 into 6, 36 numbers. The total outcome is possible by rolling dice twice in the experiment. So, out of those seven, how many times we have seen already, how many times three is appearing, how many times four is appearing, how many times two is appearing, and how many times 12 is appearing means the summation of the two outcomes. So, we have to estimate a finite probability that gives the formation of sample space of the random variables. So, it is essential to note that the experiment and the sample space used to address this question remain the same. So, we have jotted down the total probabilities of the summation of the outcome as two here, here: 2, 4, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12.

It is showing a curve. So, the highest number of times it occurs is summation 7, 1 by 6. 6 by six means 12 occurs only once, and two occurs only once. This shows the sample space distribution of these random variables 2, 12, and 7—discrete and continuous random variables.

A discrete random variable is characterized by its ability to assume, at most, a countable number of possible values, countable number. Consequently, any random variable capable of adopting either a finite number or a countable infinite number of distinct values quantifies as a discrete random variable. It is worth mentioning that there are also random variables whose set of potential values is uncountably infinite, for the continuous random variable pertains to the scenarios where the outcome of random events are numerical and yet cannot be enumerated and are infinitely divisible. So, let us see the comparison.

A discrete random variable is characterized by having possible values that are distinct points along the actual number 9. Discrete random variables are often associated with counting scenarios. A continuous random variable is defined by its possible values spanning an interval along the natural number line. Continuous random variable typically involves measurement scenarios. So, these are examples of discrete random variables.

So, the number of people in a house is always finite and is a total number, integer number 5, 3, 2, 1, 6, 7, 8, 9, 10. It is not 9.5. Number of languages a person can speak. You can speak two languages, three languages, four languages, and one language.

It is a total number. It is an integer number. Number of times a student takes a particular test before qualifying. That is also an integer—the number of collisions at an intersection, the number of intersections, and the number of accidents in the intersection.

It is also finite. Number of spelling mistakes in a document: maybe 10, maybe 5, maybe 7, maybe 11. This is a fundamental value and an integer value. For the continuous random variables, the temperature of a patient is 98.44. If the precision is double-digit, then 99.

18, 100.42 based on the precision. The height of an athlete, if I represent it in millimeters, is then submillimeter level 172.35, 160.65 like that. So, based on the scale's precision, our number of outcomes that will be coming may be infinitely divisible.

Speed of a vehicle: time taken by a person to come home from the office: 15.5 minutes, sometimes 16.5 minutes, sometimes 30.5 minutes, 30 minutes, or 31 minutes. So, there is an infinite possibility based on the precision of the time you consider.

Probability mass function. A random variable characterized by its ability to assume, at most, a countable number of potential values is termed a discrete random variable. X is a discrete random variable with n possible values, which we will label X1 and X2 up to Xn. For a discrete random variable capital X, we can define the probability of mass function Px, P of small x of x by P of small xi is equal to the likelihood of capital X is equal to xi. So, in a tabular form, the probability of X is represented by probability X1, probability X2, and probability X3 up to Xn.

So now we have to see whether these values representing the probability mass function values have some properties. So let us see. The probability mass function Px is positive, always positive or at least 0, not less than 0, at least 0, or greater than 0. So if X must assume one of the values, for example, X1, X2, and Xn, this is the possible value that Xn assumes. So, the P of xi is always greater than equal to 0, and for other values, for example, Xm is outside the ambit, so the probability of Xm will be 0 because it is not assuming that value. So now the value it believes is xi up to 1 to n; the likelihood of this summation is always 1.

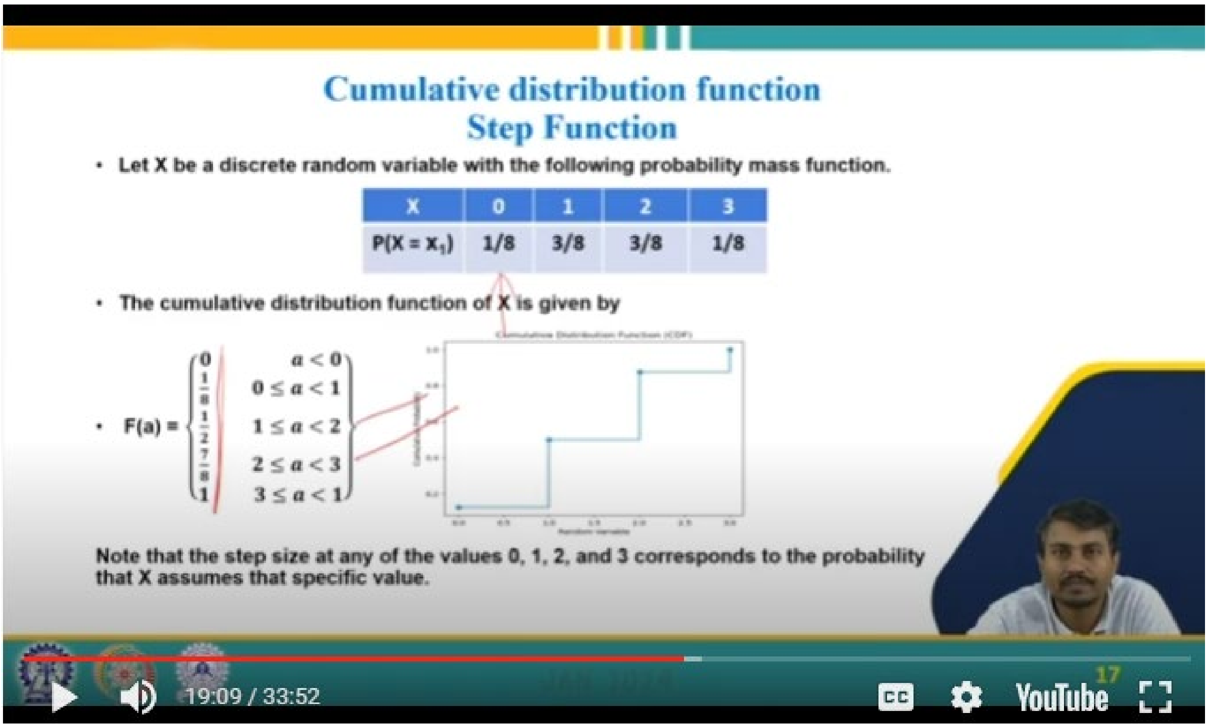
This is a fundamental property of the probability mass function. So, the P of x1 and the P of x2 up to the P of xn summation always have a probability of 1. For example, X is a random variable that assumes values of 0, 1, and 2, and all are equally likely one by 3, 1 by three probability of guessing the value 0, 1, and 3, 0, 1, and 2 is one by 3, 1 by 3, 1 by 3. So, each probability is greater than 0 or equal to 0 non-zero, and the sum of probability is 1. So, it satisfied the likelihood of a probability mass function, which is a probability mass function.

Another is flipping a coin three times; you are flipping a coin thrice. What is the sample space? Head, head, head, head, head, tail, head, tail, head, head, head, tail, tail, tail, tail, head, tail, tail, tail, tail, total 1, 2, 3, 4, 5, 6, 7, 8, 8 number of possible out. So, let us find out what the headcount is for each toss. So here is 3, here is 2, here is 2, here is 1, here is 2, here is 1, here is 1, here is 0. So, let us find out the probability mass function of the count of 8.

So 0 is only 1, 1 is coming here, 1, 2, 3, 3, 2, this is 2, 1, 2, 3, and 3 is coming is only 1. So you see here also the probability mass function properly satisfied, summation over this one by 8, 3 by 8, 3 by 8, 1 by 8 equals 1. So, what is the graph of mass functions? So, we are illustrating the random value on the x-axis and the probability value on the y-axis. That is, the probability of P xi is on the y-axis. So here, the probability mass function of outcome 0 is half probability, outcome 1 is one-fourth probability, and outcome 2 is one-fourth probability.

So, the probability mass function plot will be like this. Tossing a coin thrice means x equals the number of heads. So the number of heads we have already seen is one by 8, 3 by 8, 3 by 8, and 1 by 8 for the outcomes of 0, 1, 2, and 3. This number of times the heads will appear. So, the probability mass function graph will look like this: Just like a normal distribution if we represent that in a continuous form.

Now, cumulative distribution function. The cumulative distribution function denoted as f can be represented as follows. f of a is equal to the probability of X, capital X less than equal to a. So if X is a discrete random variable whose possible values are X1, X2, X3 up to Xn where X1 is less than equal to X2 less than equal to X3 and so on. A step function will be the distribution function of f of X. So let X be a discrete random variable with the following probability mass function; the cumulative distribution function is this: for less than equal to 0, it is 0. For less than 0, less than equal to 0, and less than one, this is only one by 8.



Less than equal to 1 less than two it is one by eight plus three by 8, 4 by eight that, is half less than equal to 2 less than 3, sorry greater than equal to 2 less than equal to 3 less than 3 for the same earlier also less than two greater than equal to 1. So this is 4 and 3, 7 by 8. For more fabulous than a is more significant than equal to or greater than equal to 3. So this is one because summing over all.

So the probability is this. So, total probability less than equal to or greater than equal to 3 equals 1. So, this can be represented like this step function for the cumulative distributions, which we can also represent by a step function. So, we must note that the step size at 0, 1, 2, and 3 corresponds to the probability that X assumes that specific value. Let us discuss a fundamental concept, which is the expectation of random variables, and this expectation concept is essential. There was much debate in the 17th century about the average outcome of the experiment. Different French mathematicians work on this problem. Notably, we have to note Pascal's contribution.

He gave the first insight and fundamental foundation of this theory. Later, Hoogians also worked on that theory and landed on the same foundation. Still, Hoogians have developed that for multiple outcomes or multiple kinds, multiple random variables. The debate was when the game continued between the two, which still needed to be finished. Who was the winner to settle that game fairly? It was a necessity to take the help of the expectation. So, this is the origin of the concept of coming out of this expectation, and nowadays, this expectation is a fundamental concept in probability.

$$\mathbf{E(X)} = \sum_{i=1}^{\infty} x_i P(x_i)$$

So let us see. So let X be a discrete random variable taking values of X1, X2, and Xn and the expected value of X denoted by E of X. E is a capital, X is also capital and referred to as the expectation of X by this capital EX is equal to the summation of i is equal to 1 to infinity Xi P Xi this is the summation. So X1 P X1, X2 P X2, X3 P X3, and so on if it is infinite. So, the expectation of a random variable can be interpreted as a long-run average, which is what the intention is. The most likely outcome is the values of the random variable across repeated independent observations.

Expectation examples. So, consider X a discrete random variable with the following probability mass function. We have already seen the outcome of the head number of times the head will appear if the coin is flipped thrice. So this is the probability. So what is the expectation? The expectation is 0 into one by 8, 1 into three by 8, 2 into three by 8, 3 into three by one by 8, and finally, it is two by 3.

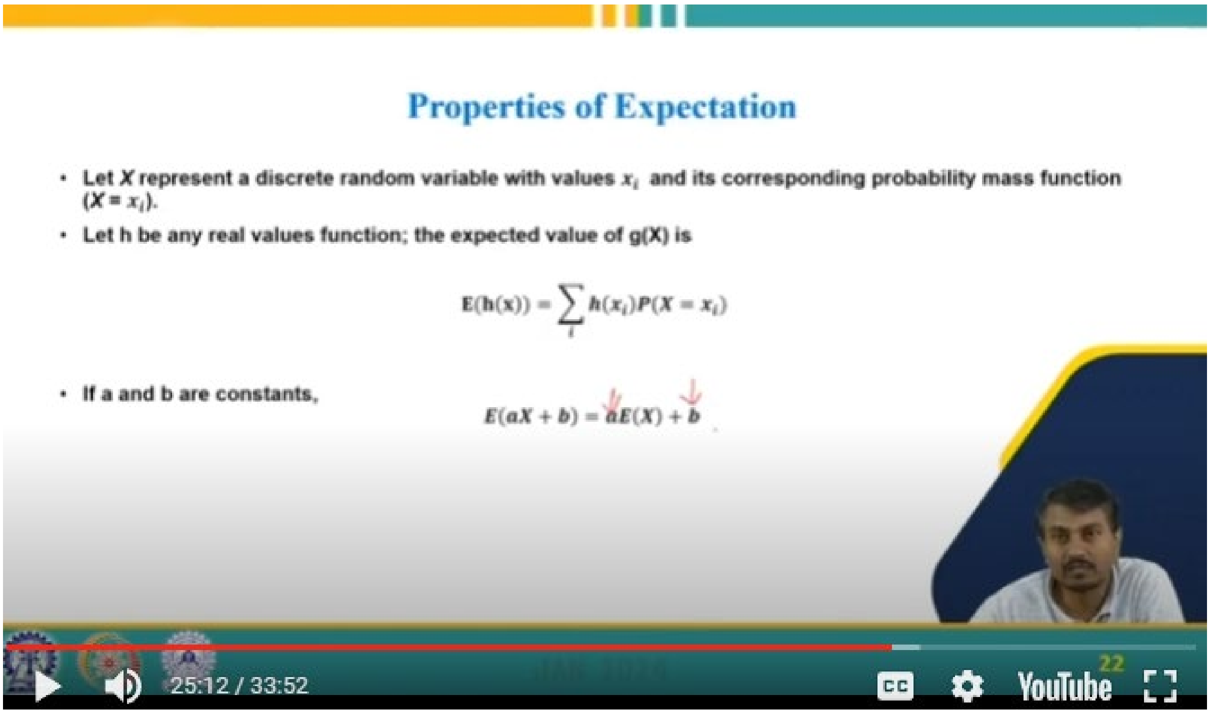
Bernoulli random variable. This is also another important concept. A random variable that assumes 1 or 0 is called a Bernoulli random variable. So, let X represent a Bernoulli random variable that assumes the value 1 with the probability of P. So, the probability distribution of this random variable is as follows. So the value one probability is P, so value 0 is one minus P.

I so expected the value of the Bernoulli random variables. Expectation EX is equal to Xi PXi summing over i is equal to 1 to infinity. So, the error is only 2. So 0 into one by one minus P is equal to 0, and 1 into P is equal to P. So, the expectation of the Bernoulli random variable is the probability of upcoming 1 P.

Discrete uniform random variables. So, let X denote a random variable that is equally probable to assume any values 1, 2, 3, and n. So, the probability mass function means that one is coming one by n, and two is coming one by n, so all are equally likely. For that, the expectation is nothing but one by n, one by n, one by n, 2, 1 by n, 2, 1, 2 into one by n, three into, so 1 plus 2 plus 3 up to n divided by n is the expectation. So that is n plus one by 2.

So, properties of expectation. So, let X represent a discrete random variable with variable values Xi and its corresponding probability mass function of capital X equals Xi. So let H be any actual value function; the expected value of G of X is equal to the expectation HX is equal to HXi, and the probability of X is equal to Xi, summing over i. So here, if A and B are constant, the expectation of E of A XB, so the constant will be multiplied here, and for the continuous expectation remains the same, that is constant itself B. But we must remember that the expectation of X square is not equal to that of X square, the expectation of the sum of two random variables. So, it is a fundamental property that the expected value of the sum of random variables equals the sum of individual predicted values.

$$E(\sum_{i=1}^{k} X_i) = \sum_{i=1}^{k} E(X_i)$$



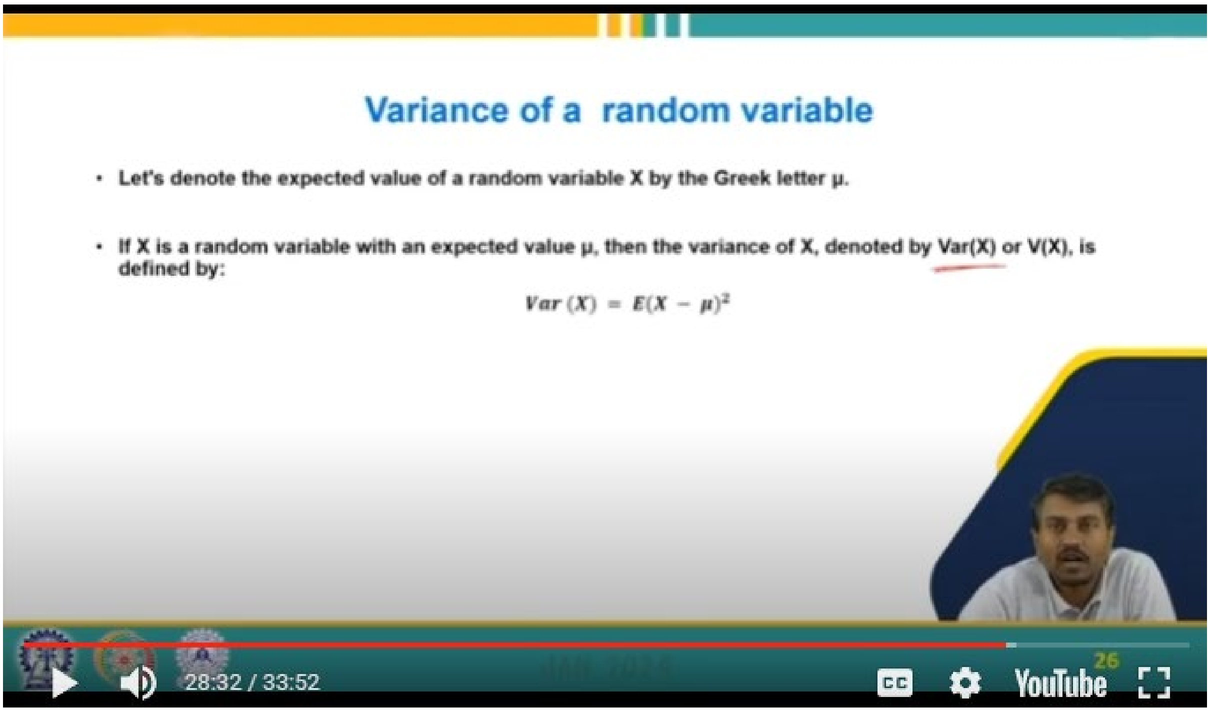
So, in other words, let X and Y be two random variables. So, the expectation of X plus Y equals the expectation of X and Y. Expectation of the sum of many random variables. So the earlier theorem expectation of E of X plus Y is equal to the expectation of X expectation of Y. So this can also be extended over many random variables. So if the number of variables is up to X1, X2, X3, Xn up, the expectation of summation is equal to the individual expectation of X1, the expectation of X2 up to Xn.

$$\mathbf{E(X+Y) = E(X) + E(Y)}$$

So, let us discuss the need for variance. Why is it needed? The expected value of a random variable provides a weighted average of its potential value, but it does not provide information about the variation or spread of these values. So, for this example, consider the random variables X, Y, and Z with their respective values and probability. X is equal to 0 with probability 1. Y is equal to minus 3 with a probability of 1 by 3.

Y equals a probability of 3 with a probability of 1 by 3. Z with a probability of minus 500, probability is half with plus 500, probability is half. So, the expectation of E X, E Y, and E Z equals 0. Here, it should be one by 2 and 1 by 2. However, the spread of Z is greater than that of Y, and the Y spread is more significant than that of X.

So, variance of random variables. So, let us denote the expected value of the random variable capital X by the Greek letter mu. So if X is a random variable with expected value mu, then the variance of X is denoted by var X. This is an essential denomination of var X. So or V X generally, we will write var X for the variance of X. So var X is equal to the expectation of X minus mu square.



Mu is also expectation, remember, of X, X minus expectation square. So, the variance of a random variable X quantifies the squared difference between the random variables and its mean mu on average. So, how do you compute? So X minus mu whole square is equal to X square mu square minus two mu capital X. So using the property of expectation, the expectation of X minus mu square, the expectation of X square plus the expectation of mu square minus two mu expectation of X.

So, the expectation of X square mu square is a constant. So that is constant mu square minus two mu expectation of X. So constant is coming here. So e X is also mu, so minus two mu square. So finally, e X square minus mu square is the variance. So variance equals e X square minus the expectation of X whole square.

- $\boldsymbol{Var(X) = E(X - \mu)^2}$
- $\boldsymbol{(X - \mu)^2 = X^2 + \mu^2 - 2\mu X}$
- **Using properties of Expectation, We know**

$$\boldsymbol{E((X - \mu)^2) = E(X^2 + \mu^2 - 2\mu X)}$$

$$\boldsymbol{= E(X^2) + \mu^2 - 2\mu E(X)}$$

$$= E(X^2) + \mu^2 - 2\mu^2$$

$$= E(X^2) - \mu^2$$

$$Var(X) = E(X^2) - (E(X))^2$$

Mu square is nothing, but mu equals E X. So it is mu square, which is the whole square. So, let us see how this variance will be. So, for a rolling dice example, the outcome is 1, 2, 3, 4, 5, 6, and all are equally likely. So the sample space is 1, 2, 3, 4, 5. So, random variable X is the roll's outcome, and the probability distribution is given here.

So, the expectation is 3.5 because one into one by 6, 2 into one by 6, 3 into one by 6, and summing over that, n plus one by 2, n is equal to 6. So 6 plus 1 divided by 2, 7 by two equals 3.5. The expectation e X square is 1, 4, 9, 16, 25, and 36 for the X square.

So you estimate that it is coming to 15.167. So, the variation is the expectation of X square, so this is 15.167 minus the expectation of X whole square, so 3.

Five whole squares. So it is coming to 2.917 variance. Bernoulli is a random variable that can assume either the value 1 or 0 is referred to as a Bernoulli random variable; we have already discussed that. So, let X be the Bernoulli random variable that takes value 1 with probability 1. So this table is known to you now. So the expectation for this X is equal to P, and the expectation of X square is equal to remains the same P. So the variance of the Bernoulli random variable is X square, which is equal to P, and this is P minus P square, so this is nothing but P into one by P.

Discrete uniform random variable, so let X represent a random variable with equal likelihood of assuming the value 1, 2, 3 up to N. So, for the probability mass function, we have already seen the expectation for this N plus one by 2, and the expectation square or expectation of X square is this N plus one into 2 N plus one by 6. So, the variance is coming from N square minus one by 12. So, for different random variables, variance is different.

- **$E(X) = (1 \times 1/n) + (2 \times 1/n) + \dots + (n \times 1/n) = (n+1)/2$**
- **$E(X^2) = (1 \times 1/n) + (4 \times 1/n) + \dots + (n^2 \times 1/n) = (n+1)(2n+1)/6$**
- **$Var(X) = E(X^2) - (E(X))^2 = (n^2-1)/12$**

So this is the reference. So, in this lesson, we have discussed the following. We have discussed the random variables, then two types of random variables we have discussed, discrete and random variables, and we have discussed the probability mass function with examples; we have discussed the cumulative distribution functions with the step functions, we have discussed the expectation and the variance and its property. Thank you.