Essentials of Data Science with R Software - 1

Professor Shalabh
Department of Mathematics & Statistics
Indian Institute of Technology Kanpur

Lecture No. 30

Degenerate and Discrete Uniform Distributions

Hello friends, welcome to the course Essentials of Data Science with R software and now, from this lecture, we are going to start our new topic that is about probability mass functions. Well, you already have done what is called a probability mass function. But now I am trying to make you learn some other things, I would like that you learn something more, you see what is the probability mass function that is a function which is trying to describe the phenomena in terms of probabilities and probability mass function is useful when the random variable is discrete, and if a random variable is continuous, then we try to use the concept of probability density function.

So, if you try to see, we already have understood that what is the meaning of probability distribution and how a random variable describes the probability over the entire range of X in terms of the probability which are taken by different values of X either at points or in intervals. Now, every phenomena can be represented by a mathematical function that is our basic assumption and that function can always be converted into a probability mass function or a probability density function, you simply have to just transform it such that it satisfies the properties of being pdf or pmf.

Now, the question is different types of phenomena can have different types of functions, which are going to help us in the competition of probability. So, the question is, if we can have some particular type of functions, which will be representing the probabilities under a specific type of conditions, do not you think that it is going to help us, that if I can match that the real phenomena is matching with this function and then I know that function and then based on the properties of that function, I can compute my probabilities and other characteristics of the data.

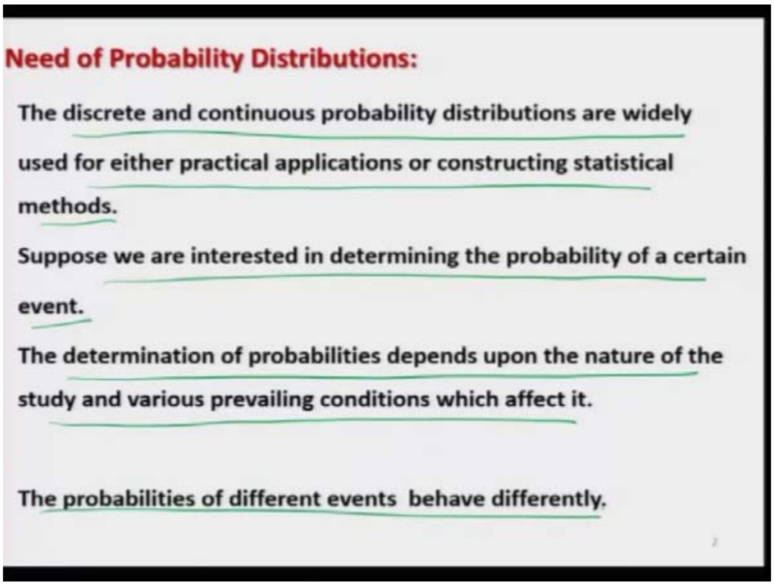
With this objective, we have different types of probability functions and each of the probability function is trying to describe the distribution of probability under a specific type of conditions. So, now what we try to do that first we try to assume that X is a discrete random variable and we try to understand some popular probability mass functions and after that, we will also consider some popular probability density functions when X is continuous.

So, from this lecture, I will try to take up the probability mass function and I will also show you that how those things can be implemented in the R software. So in this lecture, we are going to talk about two types of probability mass function, one is Degenerate Distributions and another is Discrete Uniform Distribution, well when we are talking of the uniform distribution, then this is defined for both when X is discrete or continuous. So, in this case we are going to concentrate on the discrete random variable and later on we will consider the continuous uniform distribution also.

So, now the first thing comes here, what is a degenerate distribution or degenerate probability mass function, you see the most simple thing what can occur is that the probability of each and every event is the same, just constant and the entire probability is concentrated only at one point. In that case, what will happen? That we need a mathematical function which can describe such a phenomena. So, this degenerate distribution describes these types of phenomena.

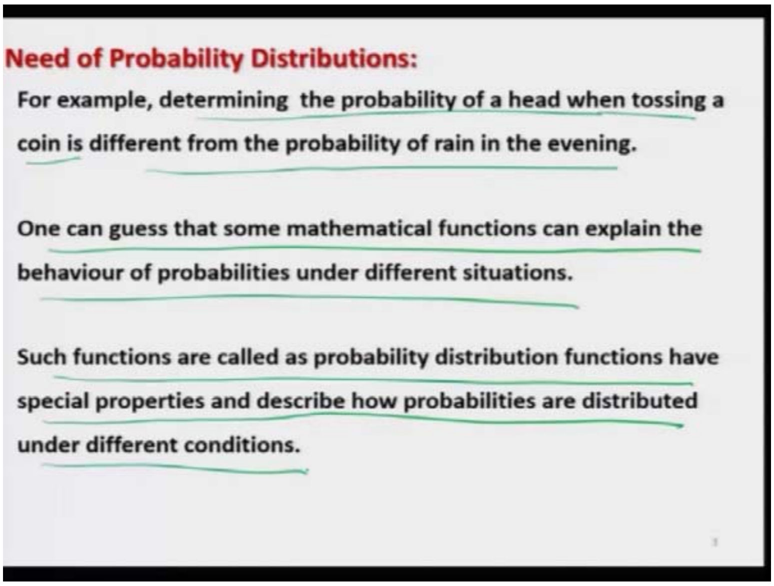
Well, I am trying to begin with a very simple one. Well, in this case actually, there is hardly any concept of probability is going to be used, but as soon as I come to the discrete uniform you will see that the things are moving and things are changing. So, let us try to begin the lecture. The first question comes, why do we need the probability distributions?

(Refer Slide Time: 05:11)



So, the first question comes here, why do we need the probability distributions, we know that the discrete and continuous probability distributions are widely used for either practical application or for constructing different types of statistical methods and suppose, we are interested in determining the probability of certain type of event, we know the phenomena, we know the nature of the phenomena that how the things are happening. The determination of probability depends on the nature of the study and various prevailing conditions which affect the phenomena. The probabilities of different events behave differently.

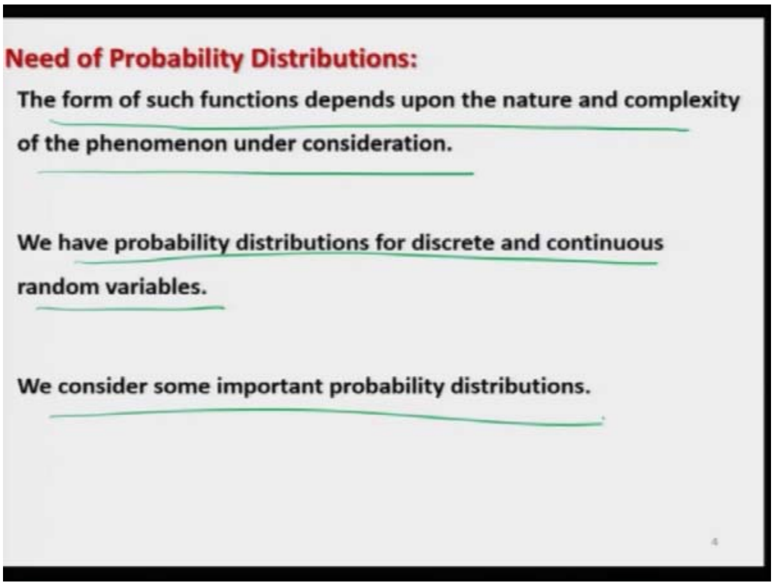
(Refer Slide Time: 06:03)



So, for example the probability of determining the head when tossing a coin is very different than determining the probability of rain in the evening these two are different events and do you think that if I asked you to compute these two probabilities without using any mathematical function, do you think that you are going to compute them in the same way? So, do not you think that it will be nice if you have got two mathematical functions which are describing you that how to compute these probabilities.

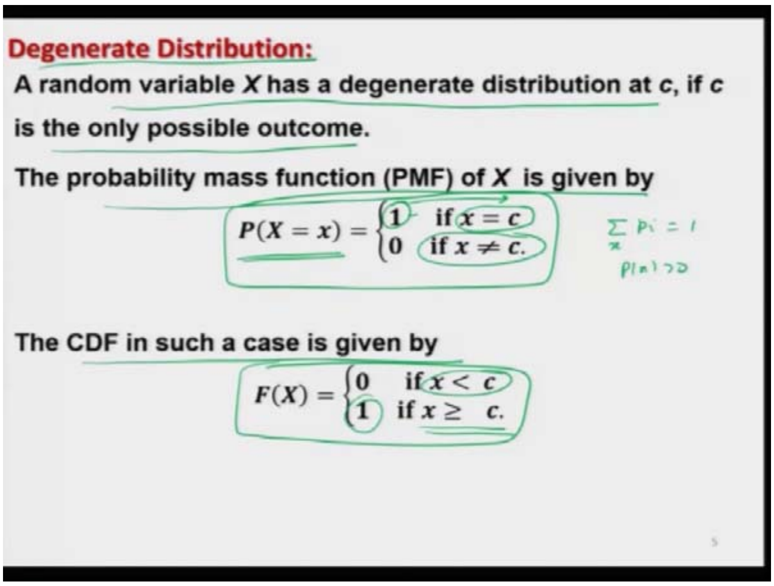
And without doing any type of thought process, you can simply use those functions to compute different types of probabilities either for tossing a coin or for probability of rain in the evening. So, one can actually guess some mathematical functions can explain behaviour of probabilities under different situations and such functions are called as probability distribution functions and have special properties and describe how probabilities are distributed under different types of conditions.

(Refer Slide Time: 07:11)



So, the form of such function depends upon the nature and complexity of the phenomena under consideration, for example and the first probability function which I am going to consider here you can see that it is one of the most simple probability functions and as soon as you complicate the phenomena, the form of the mathematical function will also become more complex. So, we have the probability distribution for the discrete as well as continuous random variable. So, now we are going to consider some important probability distributions, which are useful for us.

(Refer Slide Time: 07:51)



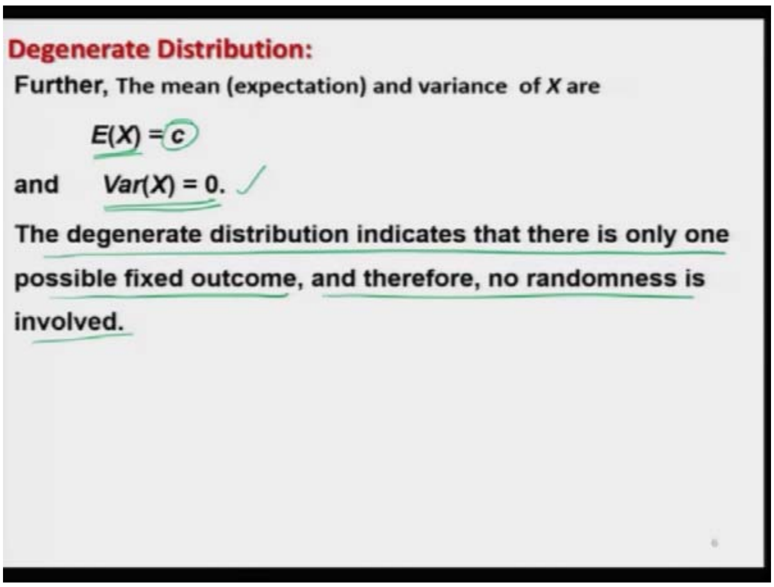
So, first we try to take here the degenerate distribution. So, now if you try to see, if there is any phenomena, where the entire probability of the event is concentrated only at one value that means, if the random variable takes that value, the probability is 1 otherwise it is 0. These types of situations need to have a particular form of mathematical function which can describe such behaviour.

And if you try to see, if you try to describe such behaviour through a mathematical function, do not you think that I can write down here as a $P(X = x) = \begin{cases} 1 & \text{if } x = c \\ 0 & \text{if } x \neq c. \end{cases}$, means if x is smaller than c is greater than c, whatever it is, it is always taking the value 0, that means the entire probability is concentrated only at the point x equal to c.

So, now do not you think that this is trying to describe the behaviour of the phenomena? And now, so, you can think about is such a function and in such a case, we say that a random variable X has a degenerate distribution at c, if c is the only possible outcome and the probability mass function or the pmf of x is given by this function probability of x equal to x which takes value 1 if X is equal to c, otherwise it is 0 and you can actually check here that both the condition of a function to be probability mass function that summation over x and probability of x to be greater than 0, they are satisfied.

And in case if you try to find out the cumulative distribution function, CDF in such a cases, this can be found very easily as $F(X) = \begin{cases} 0 & \text{if } x < c \\ 1 & \text{if } x \geq c. \end{cases}$. So, now you can see here that whenever in practice you have this type of phenomenon, you can directly compute any type of probability using this function.

(Refer Slide Time: 10:35)

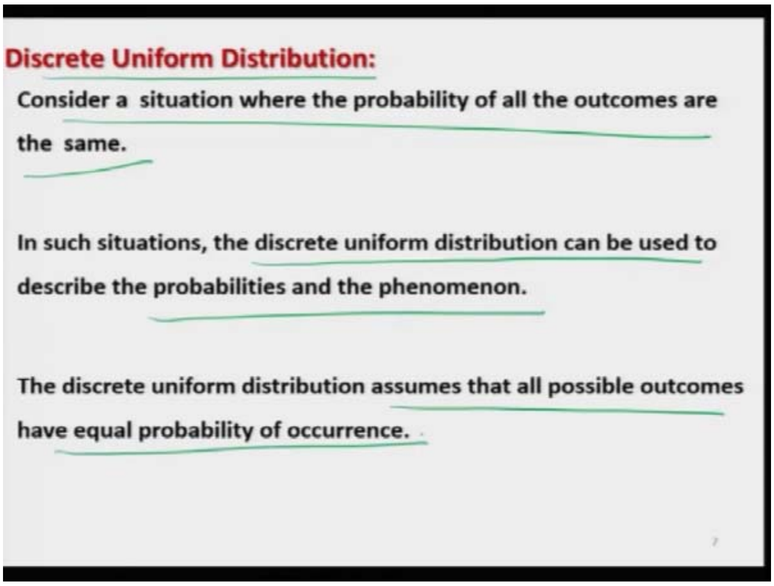


Well, these functions are not a means popular in real life, but they help us in several type of mathematical operations, when we are trying to model a and since, we are considering here a random variable, so that means, this random variable will have certain properties and those properties can be described in terms of mean, variance, skewness, kurtosis etc. or different types of moments I would say, in general.

So, in this case if you try to find out the mean and variance of x that is the expected value of X, this will come out to be here c that is very straightforward to find out and variance of X will come out to be 0 that means, that there is no variability means at all the points it is taking value 0 otherwise it is taking only one value at X equal to c.

And in that case the entire probability is concentrated only at one point, so there is practically no randomness that is what is indicating by this expression variance of X equal to 0. So, this degenerate distribution indicate that there is only one possible fixed outcome and therefore, there is no randomness involved. You can see here this is also indicating the same thing what do we expect?

(Refer Slide Time: 12:01)

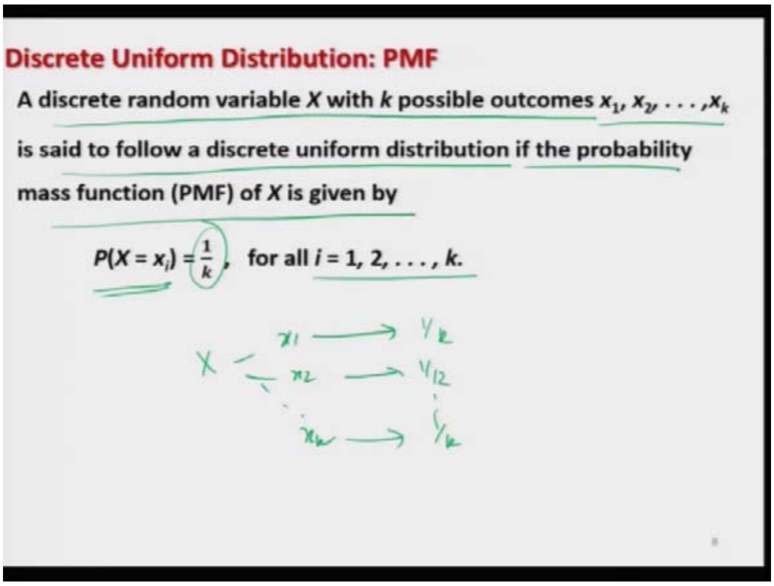


So, now I try to come to one more distribution which is called as discrete uniform distribution. So, we try to consider a situation where the probability of all the outcomes are the same. Can you think about such a situation where the probabilities of all the outcomes are going to be the same, why you have done it, tossing a coin, rolling a dice? In case of rolling a dice, the probability was just 1 by 6, either X takes only 1, 2, 3, 4, 5 or 6. In that toss of a coin, the probability was always 1 by 2, either it is head or tail.

So, any phenomenon in which the probability of occurrence of an event remains the same, how to describe this type of probability that is the question. So, we need here a mathematical framework which can describe this type of thing and then do not you think that this is practically possible, whenever you are trying to draw any ball from a box, just random without looking into the box, the probability of all the balls is the same. So, how to describe that type of phenomena and in this case, we try to take the help of discrete uniform distribution.

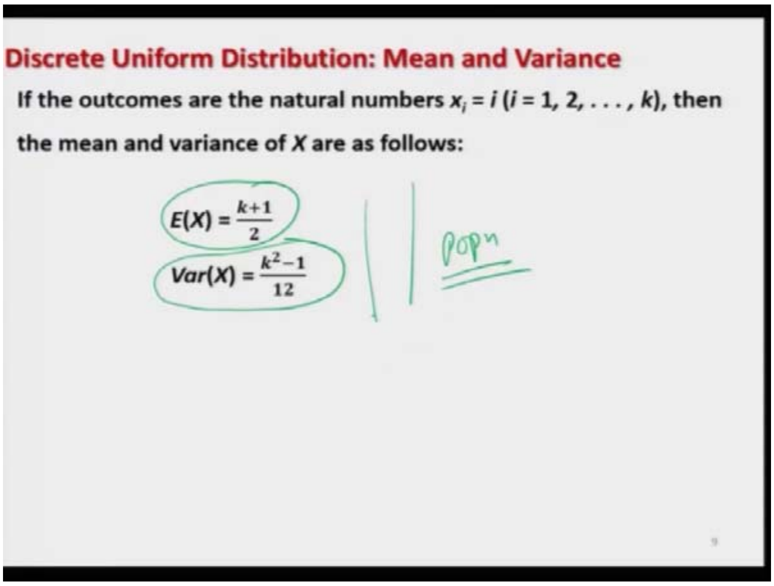
So, in such situations, wherever the probability of all the outcomes are going to be the same, we use the discrete uniform distribution which describes the probabilities and phenomena and this distribution assumes that all possible outcomes have equal probability of occurrence.

(Refer Slide Time: 14:10)



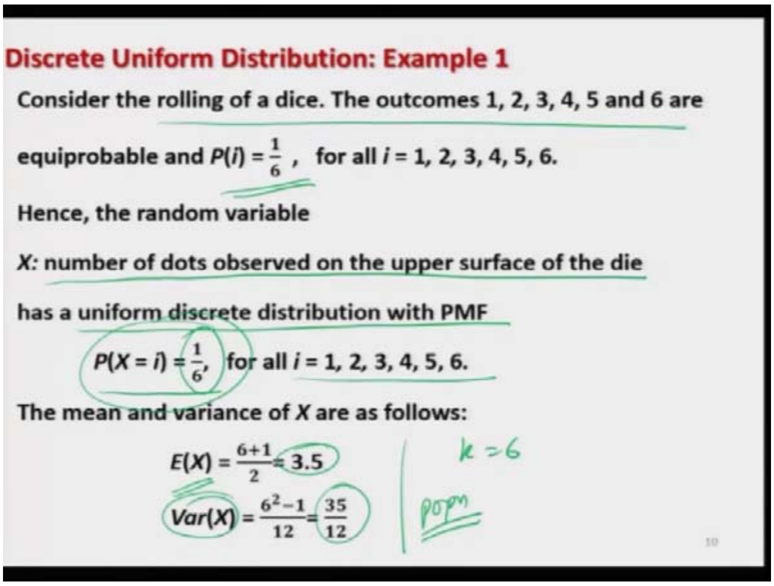
So, now the question is how are you going to describe such phenomena through a mathematical function and that is going to be converted into a discrete uniform distribution. So, now I can say that discrete random variable X with k possible outcomes say $x_1, x_2, \ldots, x_k$ is said to follow a discrete uniform distribution, if the probability mass function of x is given by $P(X = x_i) = \frac{1}{k}$, for all $i = 1, 2, \ldots, k$ that means, x takes value x_1 or x_2 or say here x_k, the probability of occurrence of the event is always 1 upon k, 1 upon k, up to here, 1 upon k. So, you can see here that this mathematical function is very well trying to define the phenomena in a mathematical way.

(Refer Slide Time: 15:05)



And in case if you try to find out the mean and variance of this random variable, then you can very easily find them as expected value of X will come out to be (k + 1)/2 and variance of X will come out to be $(k^2 - 1)/12$. Well, there is a issue that what is the meaning of this mean and variance and because these values are going to hold in the entire population, but practically you will be working in a sample. So, that is what is my objective that I will try to take the help of R software and we will try to show you that what is the meaning of this thing.

(Refer Slide Time: 15:54)



So, now, let us try to take a very simple example and try to see what is the meaning of this thing. So, consider the rolling of a dice there will be six possible outcomes the numbers 1, 2, 3, 4, 5 and 6 which are occurring with the probability 1 by 6. So, hence this these outcomes are the outcome of a random variable.

So, in case if I try to define here let X be the number of dots observed on the upper surface of the dice and so, this can have a uniform discrete distribution with a probability mass function probability that X equal to i is simply 1 upon 6 for all i goes from 1, 2, 3, 4, 5, 6 and in this case, if you try to find out here the mean and variance you already have found these expression which are $(k + 1)/2$ and $(k^2 - 1)/12$, if you try to substitute here k equal to 6 you can get here expected value of X here as say 3.5 and variance of X will come out to be 35 upon 12.

Now, what is the meaning of this thing I will try to show you because these are the mean and variance for the population. So that means, if you try to repeat the experiment for a large number of time then on an average the mean and variance will come out to be close to 3.5 and 35 by 12 and you also know what is the meaning of the probability of X equal to i is 1 upon 6 that we already have done many times.

So, now let us try to conduct these things in R software. But before that, I would like that you try to first understand this concept and whatever mean and variance I have given you here, try to find them yourself do some mathematical algebra. Well, my objective is not to give you the entire algebra as I promised in the beginning, but it is very important for you to understand how these expressions are coming and these are very simple expressions.

Once you do it, you will get confidence and try to think about it what is the meaning of expected value of X and variance of X and how you can show that X is a random variable which is following this probability distribution function or probability mass function. So, I will say why do not we take today a break, you try to think about them and I will see you in the next lecture where I will try to show you how you can apply these things on the R software and then possibly I will try to repeat the same thing in a very different concept.

So, in that next turn, we are going to look at the same thing for the discrete uniform distribution in a different context. You try to practice it and I will see you in the next lecture. Till then, goodbye.