1. Introduction

Welcome to the series of E-Learning session on Probability Distribution. Today, we are going to discuss on a very important concept known as Continuous Random Variable.

At the end of this session, you will be able to understand and explain:

- What is a Random Variable
- What are the different kinds of Random Variable
- What is Continuous Random Variable
- How to calculate Probability of Continuous Random Variable
- What are the properties of Continuous Random Variable
- What is a probability density function
- What are the properties of probability density function
- How to calculate probability of continuous random variable using probability density function
- Definition and properties of distribution function

Do you have any idea what is a Random Variable? Before moving forward to the Random Variable let us understand the difference between a variable and a constant.

A variable is a one that changes its value depending on time, place, person etc, otherwise it is called constant.

Whereas, constant is one which never changes its values at any situation.

For example: Let us take a person who wishes to travel from location A to location B.

Assuming that there is only one route, the distance between these 2 locations will be same irrespective of time, speed, mode of transport etc.

Whereas the time taken to reach from location A to location B, which is variable depends on the speed of the vehicle in which he is travelling.

If the speed is high, time taken is less and if the speed is low it will take more time.

Thus this example clearly shows that, distance between two locations is a constant; whereas time taken to reach from one location to the other is a variable, since it depends on speed.

Hope, you have understood the difference between Variable and a Constant.

2. Random Variable and Its Types

I am sure you are quiet familiar with the word Random Experiment and Sample Space.

Consider a random experiment of tossing 2 unbiased coins. The sample space will be S= {HH, HT, TH, TT}.
Let X is the variable which denotes head. Then X can values 0 if TT (Tails) appears, in this case probability is 1/4 , it can value 1 with probability 2/4 in case of HT or TH and 2 with probability 1/4.
So, we are able to assign probability for each value of X, which is possible in this experiment.

You can observe that the variable X is taking real values with definite probability. Hence, X is a discrete random variable.

The other kind of random variable is continuous, which takes all possible values on the real line R (-∞ to +∞).

For example: In an interval of (25, 50), the variable number of students in different class can take only integers like 26, 28, 30, 35, 40 etc. but can't take values like 26.5, 28.3, 30.6, 35.2 etc.

If we change the variable as weight of students instead of number of students, then the variable can take values like 26.5 Kgs, 28.3 Kgs, 30.6 Kgs, 35.2 Kgs etc.

It is clear from the above example that continuous random variable can take all possible values in the given interval, whereas discrete random variable can take only limited values.

Thus, A *Random Variable* is a variable which takes value with a definite probability.

In a simple words,a *random variable*, usually written *X*, is a variable whose possible values are numerical outcomes of a random experiment.

Thus as discussed Random Variable is of two types Discrete and Continuous.

If the values of the variable are captured through counting we may say it is a discrete variable.
Where the values of the variable are captured through measurement then we may say it is a continuous variable.

Let a random experiment is conducted and let S be the continuous sample space. That is set of all possible outcomes when a random experiment is conducted.
Then define a random variable X on the sample space S. Since, X is defined on the continuous sample space, X is also continuous. Then it can be observed that X can take all possible real values on the real line from - ∞ to + ∞, depending on the sample space.

Thus, a *continuous random variable* is a function which is defined on a sample space (S), associated with a given random experiment and taking all possible values on the real line R (-∞ to +∞).

We always denote the random variable by capital letters like X, Y, Z…... and the values taken by the random variable is denoted by lower cases x, y, z….

Let f(x), is a probability function (Probability Density Function (PDF) is a statistical measure that defines a probability distribution for a random variable and is often denoted as f(x)) in associated with a continuous random variable X in a given interval say - ∞ to + ∞.

Suppose, you are interested in find the probability that random variable X taking the values " ≤ " to x. to find the probability you have to integrate the given probability function keeping the range from - ∞ to x, that is

$P(X \leq x) = \int_{-\infty}^{x} f(x)\,dx.$ **(Formula 1)**

Let f(x), is a probability function (we need to discuss about probability density function that, we will take it coming sessions) in associated with a continuous random variable X in a given interval say a to b.

Suppose, you are interested to find the probability that random variable X taking the values "a", to find the probability you have to integrate the given probability function keeping the range from a to x, that is

$P(X \leq x) = \int_{a}^{x} f(x)\,dx.$ **(Formula 2)**

Suppose, you are interested to find the probability that random variable X taking the values x to "b", to find the probability you have to integrate the given probability function keeping the range from b to x, that is

$P(X \geq x) = \int_{x}^{b} f(x)\,dx.$ **(Formula 3)**

3. Properties on Continuous Random Variable

Now, I am going to talk about properties of continuous random variable.

First property:
Like discrete random variable, continuous random variable can't take any exact value. The probability of a continuous random variable X assuming the exact value x is always zero. That is, P(X=x) = 0.

Second property:
For a continuous random variable, since the function is continuous at a point x, the probability that X assuming value at a point x is zero.
Therefore, it follows that "≤" or < does not make any difference.
i.e., $P(x \le X \le y) = P(x < X < y)$.

Let us consider one specific example,
$f(x) = Cx^2$, $0 < x < 1$, **(Formula 4)**
If f(x) has a probability density Cx square , 0<x<1,
find the probability that 1/3 < x < 1/2.

To find the probability that x lies between 1/3 and 1/2,
That we need to find
$P(1/3 \le x \le 1/2) = \int_{\frac{1}{3}}^{\frac{1}{2}} Cx^2\,dx.$ **(Formula 5)**
$= C\int_{\frac{1}{3}}^{\frac{1}{2}} x^2\,dx.$
$= C\,\frac{x^3}{3}$ between $\frac{1}{3}$ and $\frac{1}{2}$ **(Formula 6)**
= C/216
Therefore the probability that x lies between 1/3 and 1/2 , for the given function is C/216.

We are given the probability function; we need to find the probability that the random variable take the value in between 1 and 2. To solve this problem you need to know integration of exponential function. Hope, all of you know integrating exponential functions.

Let f(x) is the probability associated with the continuous random variable X.

Given $f(x)= \begin{cases} 3e^{-3x}, x>0 \\ 0, Otherwise \end{cases}$ **(Formula 7)**

Find $P(1 \le x \le 2)$

We have
$P(1 \le x \le 2) = \int_1^2 3e^{-3x}\,dx$ **(Formula 8)**

Two find required probability,
We need to find ∫evaluate between 1 and 2 3e minus 3x dx = 0.0473

=$\int_1^2 3e^{-3x}\,dx$

=3$[e^{-3x}/-3]$ evaluate between 1 and 2 **(Formula 9)**

= 3 $[(e^{-6} - e^{-3})/3]$

=$(e^{-6} - e^{-3})$

From exponential tables we get,
= 0.0473

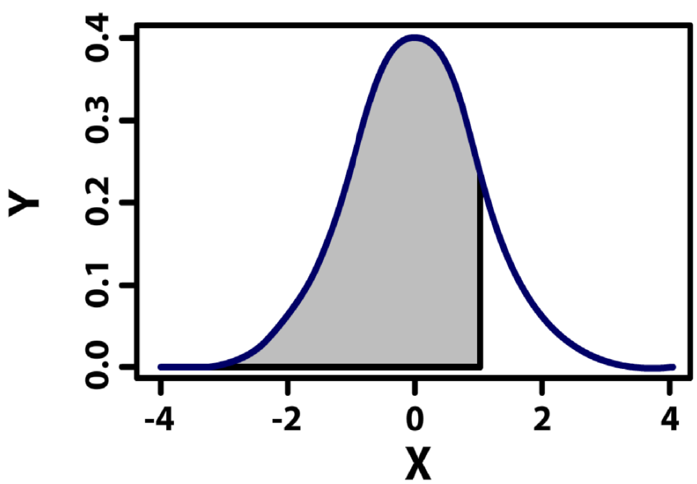
Thus, from the above slides we have learned about the Random Variable, different kinds of Random Variable, Continuous Random Variable, how to calculate Probability of Continuous Random Variable, and the properties of Continuous Random Variable.

4. Probability Density Function

Let us now discuss about a very important concept known as Probability Density Function.

As you are aware that probability mass function (P.M.F) is associated with the discrete random variable.
The probability of a continuous random variable assuming any value in the given range is calculated using the function called probability density function abbreviated as P.D.F. Probability Density Function, since it is a continuous function, the total area falling under the curve drawn will give you the probability of a random variable taking a particular value.

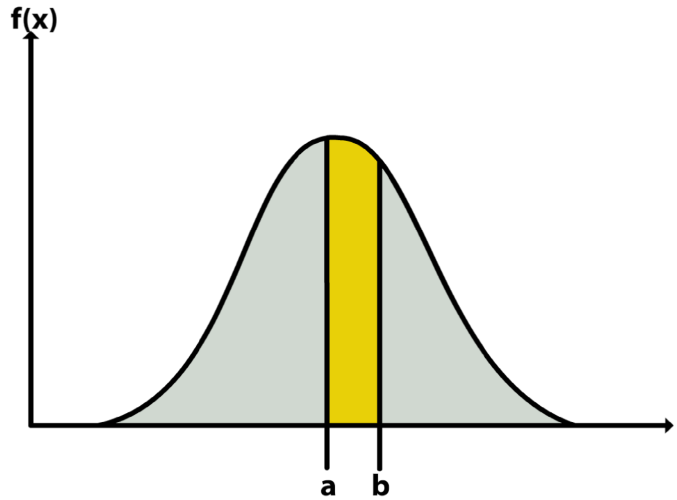
Figure 1



Let us consider a random variable X which can assume values in the interval (a, b) and f(x) denote the probability function associated with the random variable X.

Consider values taken by the variable on X axis and f(x) on Y axis.

Figure 2



If we join all the points(x, f(x)), the curve emerges will give the probability of a random variable X for different values.
Let X be the random variable and f(x) be the probability function associated with the random variable X in the range (a, b). The function f(x) has to satisfy certain conditions to call it as P.D.F.

There are 3 important properties of P.D.F. Let X be the random variable and let f(x) be the probability function associated with the random variable X in the range (a, b). Then f(x) is called P.D.F of X, if it satisfies the following conditions:

- Non – negativity. For all values of x, function f(x) should be nonnegative. That is, f(x) should not take any negative values.
  It is true even if a = - ∞ and b = ∞.
  Thus, $f(x) \geq 0$ for all values of x in the range (a, b).

- Function f(x) should always a continuous function for all values of x in the range (a, b). The function should not have any discontinuity point in the interval (a, b).
  Thus, f(x) is continuous for all values x in the range (a, b).

- Total integral should always be one or total probability is always one. That is when the function f(x) is integrated over the range (a, b), the result should always be equal to unity (1).

From these conditions we can say that all P.D.F.'s are non negative continuous functions, but converse is not true.

Thus it is clear from the properties that:

1. It is true for all values a and b on the real line - ∞ to + ∞
2. Students should always keep in mind that all P.D.F's are continuous functions, but all continuous function need not be a P.D.F
3. The continuous function becomes P.D.F it satisfies property 1 and 3 also.

So far, we have understood the definition and properties of P.D.F. Now, let us understand how to calculate probability for a give value of X using P.D.F.

To find the probability for a given value we need to integrate the function in the required range.
If f(x) is a P.D.F. associated with the continuous random variable X in a given interval (a, b) then probability of X for a give values of x(x being the maximum value) we need to find $P(X \leq x) = \int_a^x f(x)dx.$ for a$\leq x \leq$b. **(Formula 10)**

Here, you need to integrate the function f(x) by keeping left limit of the interval as lower limit of the integral and the specific value of x as the upper limit. If you solve this integration, the resultant will be the required probability.

Similarly, if you want to find the probability such that x is the minimum value, then we need to find $P(X \geq x) = \int_x^b f(x)dx.$ for a$\leq x \leq$b, **(Formula 11)**

That is you need to integrate the function f(x) by keeping specific value of x as a lower limit of the integral and right limit the as the upper limit of the integral. If you solve this integration, the resultant will be the required probability.
Another, method of finding P(X ≥ x) is using complementary properties of probability.

That is: $P(X \geq x) = 1 - P(X \leq x) = 1 - \int_x^a f(x)\, dx$ **(Formula 12)**

(Using complementary probability rule).

5. Examples

Example 1:

Now let us see how we can use P.D.F. to find probability.
Take an example of sales per day in a retail stores. That is random variable X denote the sales per day (in lakhs rupees). It has been observed that the sales per day is following pattern, based on which we have the density function as

$$f(x) = \begin{cases} \text{cx} & 0 \le x \le 2 \\ 0, & otherwise \end{cases}$$ **(Formula 13)**

We need to find the values of constant C and also we need to find the probability that on a given day a sale is Rs. 60,000.

Firstly, we need to find the constant c:
We use the property that the total probability is always one.
That is , $\int_a^b f(x)dx.$ = 1, $for\ all\ values\ of\ x$ in the range (a, b)

That is we need to evaluate $\int_0^2 cx\, \boldsymbol{d}x = 1$ **(Formula 14)**

$$\text{c}\int_0^2 x\, \boldsymbol{d}x = 1$$

c[x2/2] = 1 we need to solve this between 0 and 2
c[2^2/2 - 0^2/2] = 1, c[
2 - 0] = 1
On solving, we get c = ½ = 0.5

Example 2:

We need to check whether the given function is a P.D.F.
It can be observed that the function is non-negative for all values of x and the function is continuous.

It can also be observed that the $\int_0^1 \mathrm{f(x)dx.}$ = 1, **(Formula 15)**

Hence, it is P.D.F.
A random variable X has the P.D.F.

f(x) = {2x,0≤x≤1
0,otherwise

Find the probability of the following:
1. P(X < 0,X > 1)
2. P(X ≤ 0.5)
3. P(X ≥ 0.8)

It is important to observe that P(X< 0,X > 1) = 0, because the random variable X takes values only in the interval 0 and 1, outside this interval the variable can't take any value. That is why probability is zero.

To find $P(X \leq 0.5)$
we need to evaluate

$$P(X \leq 0.5) = \int_0^{.5} f(x)dx.$$

$$= \int_0^{0.5} 2xdx. \qquad \textbf{(Formula 16)}$$

$$= \int_0^{0.5} 2\ x\ dx.$$

$$= 2\int_0^{0.5} xdx.$$

= $2[X^2/2]$ we need to solve this between 0 and 0.5
= $2[0.5^2/2 - .0^2/2]$
= 2[0.25 - 0] = 0.5
Therefore, $P(X \leq 0.5) = 0.25$

Let us understand one function associated with P.D.F. called distribution function abbreviated as D.F.
In the given interval, D.F. gives the probability of values assumed by the variable till a specific value x.
If f(x) is a P.D.F. associated with the continuous random variable X in a given interval $(-\infty, \infty)$ then distribution function of X denoted by F(x) is defined as:

1. $F(x) = P(X \leq x) = \int f(x)dx$ for $-\infty \leq x \leq \infty$
2. We write distribution function as D.F.
3. It is customary to denote P.D.F. as f(x) and D.F. as F(x).

Following are the very important properties of D.F.
For any values of a and b, $a < b$, it is always true that
$P(a < x \leq b) = F(b) - F(a)$
Another important property is that D.F. is always monotonically non decreasing function.
For any given value of a and b, $a < b, F(a) \leq F(b)$
Since, D.F. is also a probability function it always lies between 0 and 1.

Here's a summary of our learning in this session:

- What is a Random Variable
- What are the different kinds of Random Variable
- What is Continuous Random Variable
- How to calculate Probability of Continuous Random Variable
- What are the properties of Continuous Random Variable
- What is Probability Density Function
- Properties of P.D.F.
- How to calculate probability for a continuous random variable using P.D.F.
- Application of P.D.F.
- Distribution function and their properties.