Random Variables

Random Variables are basically a function that maps from the set of sample spaces to a set of real numbers. Random variables are defined over a sample space of any random experiment. Values of random variables correspond to the outcomes of the random experiment.

What is Random Variable?

Random variable in statistics is a mathematical concept that assigns numerical values to outcomes of a sample space. There are two types of Random Variables, Discrete and Continuous. A random variable is considered a discrete random variable when it takes specific, or distinct values within an interval. Conversely, if it takes a continuous range of values, then it is classified as a continuous random variable. Random variables are generally represented by capital letters like X and Y. This is explained by the example,

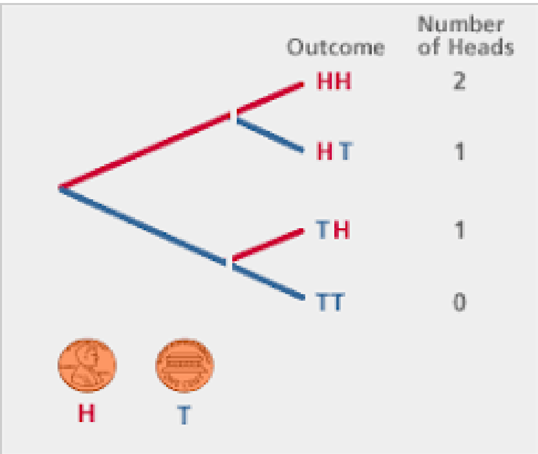
Example: If two unbiased coins are tossed then find the random variable associated to that event.

Solution:

Suppose Two (unbiased) coins are tossed

X = number of heads. [X is a random variable or function]

Here, the sample space S = {HH, HT, TH, TT}



Random Variable Definition

We define random variable a function which maps from sample space of an experiment to the real numbers. Mathematically, Random Variable is expressed as,

$$\mathbf{X: S \rightarrow R}$$

Where,

- **X** is Random Variable (It is usually denoted using capital letter)
- **S** is Sample Space
- **R** is Set of Real Numbers

Suppose a random variable X takes m different values i.e. sample space

$X = \{x_1, x_2, x_3 \ldots\ldots\ldots x_m\}$ with probabilities

$P(X = x_i) = pi$

where $1 \leq i \leq m$

The probabilities must satisfy the following conditions:

- $0 \leq p_i \leq 1$; where $1 \leq i \leq m$
- $p_1 + p_2 + p_3 + \ldots\ldots. + p_m = 1$
- Or we can say $0 \leq p_i \leq 1$ and $\sum p_i = 1$

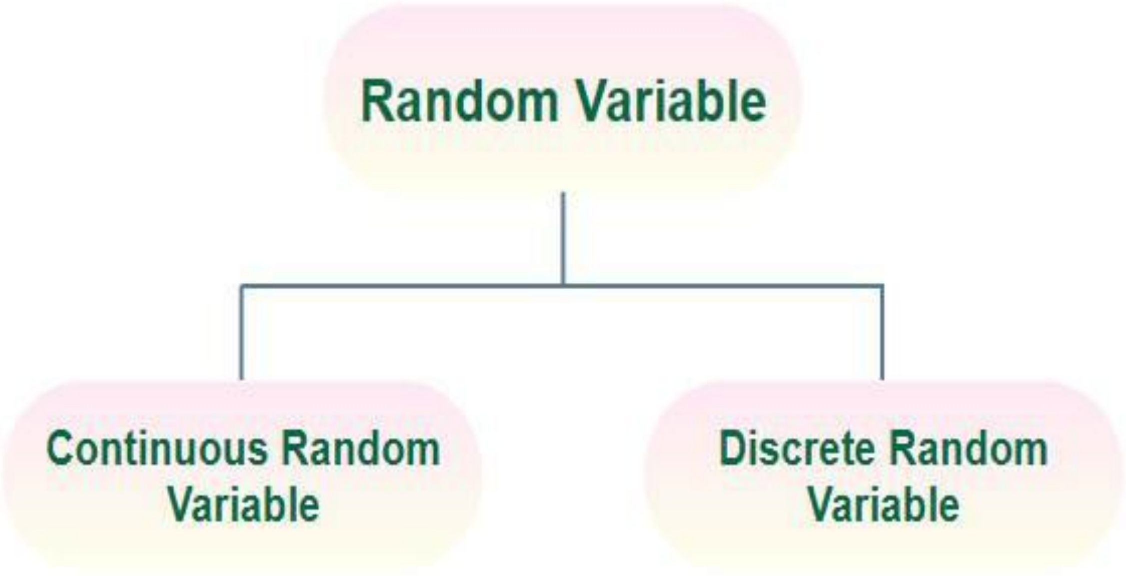
For example:

Suppose a dice is thrown (X = outcome of the dice). Here, the sample space S = {1, 2, 3, 4, 5, 6}. The output of the function will be:

- $P(X=1) = 1/6$
- $P(X=2) = 1/6$
- $P(X=3) = 1/6$
- $P(X=4) = 1/6$
- $P(X=5) = 1/6$
- $P(X=6) = 1/6$

There are two basic types of random variables,

1. Discrete Random Variables
2. Continuous Random Variables



➢ Discrete Random Variable

A random variable X is said to be discrete if it takes on a finite number of values. The probability function associated with it is said to be

PMF = Probability Mass Function P(xi), if

- $0 \leq pi \leq 1$
- $\sum pi = 1$ where the sum is taken over all possible values of x

➢ Discrete Random Variables Example

Example: Let S = {0, 1, 2}. Find the value of P (X = 0).

x_i	0	1	2
Pi(X = xi)	P_1	0.3	0.5

Solution:

We know that the sum of all probabilities is equal to 1. And P (X = 0) be P1

P1 + 0.3 + 0.5 = 1

P1 = 0.2

Then, P (X = 0) is 0.2

- **Continuous Random Variable**

A random variable X is said to be continuous if it takes on an infinite number of values. The probability function associated with it is said to be PDF (Probability Density Function).

PDF (Probability Density Function)

If X is a continuous random variable. $P(x < X < x + dx) = f(x)dx$ then,

- $0 \le f(x) \le 1$; for all x
- $\int f(x)\, dx = 1$ over all values of x

Then P (X) is said to be PDF of the distribution.

Random Variable Formula:

There are two main random variable formulas,

- **Mean of Random Variable**
- **Variance of Random Variable**

Let's learn about the same in detail,

Mean of Random Variable:

For any random variable X where P is its respective probability, we define its mean as,

$$\textbf{Mean}(\mu) = \sum X.P$$

where,

- **X** is the random variable that consist of all possible values.
- **P** is the probability of respective variables

Variance of Random Variable:

The variance of random variable tells us how the random variable is spread about the mean value of the random variable. Variance of Random Variable is calculated using the formula,

$$\mathbf{Var(x) = \sigma^2 = E(X^2) - \mu^2}$$

where,

$$E(X^2) = \sum_{i=1}^{n} (x_i)^2 p_i$$

and

$$E(X) = \sum_{i=1}^{n} x_i p_i$$

Mean And Variance Of Random Variable

x	P(x)	x^2	x^2 * P(x)
1	0.10	1*1 = 1	1 * 0.10 = 0.10
2	0.30	2*2 = 4	4 * 0.30 = 1.20
3	0.45	3*3 = 9	9 * 0.45 = 4.05
4	0.15	4*4 = 16	16 * 0.15 = 2.40

Mean Formula:

$$\mu_x = \Sigma[x * P(x)]$$

$$\mu_x = 2.65$$

Variance Formula:

$$\sigma^2{}_x = \Sigma[x^2 * P(x)] - \mu_x^2$$

Random Variable Functions:

For any random variable X if it assume the values $x_1, x_2, \ldots x_n$ where the probability corresponding to each random variable is $P(x_1)$, $P(x_2), \ldots P(x_n)$, then the expected value of the variable is,

Expectation of X: $E(x) = \sum x.P(x)$

Now for any new random variable Y in which the random variable X is its input, i.e. Y = f(X), then cumulative distribution function of Y is,

$$\mathbf{F_y(Y) = P(g(X) \leq y)}$$

Probability Distribution and Random Variable:

- **For a random variable its probability distribution is calculated using three methods,**

A. Theoretical listing of outcomes and probabilities of the outcomes.
B. Experimental listing of outcomes followed with their observed relative frequencies.
C. Subjective listing of outcomes followed with their subjective probabilities.

Probability of a random variable X that takes values x is defined using a probability function of X that is denoted by f (x) = f (X = x).

There are various probability distributions that are,

- **Binomial Distribution**
- **Poisson Distribution**
- **Bernoulli's Distribution**
- **Exponential Distribution**
- **Normal Distribution**

Random Variable Example

Example 1: Find the mean value for the continuous random variable, $f(x) = X^2$, $1 \leq x \leq 3$

Solution:

Given,

$f(x) = x^2$

$1 \leq x \leq 3$

$E(x) = \int_1^3 x.f(x)dx$

$E(x) = \int_1^3 x.x^2.dx$

$E(x) = \int_1^3 x^3.dx$

$E(x) = [x^4/4]_1^3$

$E(x) = 1/4\{3^4 - 1^4\} = 1/4\{81 - 1\}$

$E(x) = 1/4\{80\} = 20$

<u>Example 2:</u> Find the mean value for the continuous random variable, $f(x) = e^x$, $1 \leq x \leq 3$

<u>Solution:</u>

Given,

$f(x) = e^x$

$1 \leq x \leq 3$

$E(x) = \int_1^3 x.f(x)dx$

$E(x) = \int_1^3 x.e^x.dx$

$E(x) = [x.e^x - e^x]^3_1$

$E(x) = [e^x(x - 1)]^3_1$

$E(x) = e^3(2) - e(0)$

$E(x) = 2e^3$

$$\int \underline{x}\, \underline{e^x dx}$$

$$\int \underline{u}\, \underline{dv} = uv - \int v\, du$$

$$= xe^x - \int e^x\, dx$$

$$= xe^x - e^x .$$

$u = x$

$dv = e^x dx$

$du = dx$

$v = e^x$

Random Numbers in Python

Python defines a set of functions that are used to generate or manipulate random numbers through the random module.

Functions in the random module rely on a pseudo-random number generator function random(), which generates a random float number between 0.0 and 1.0. These particular type of functions is used in a lot of games, lotteries, or any application requiring a random number generation.

Let us see an example of generating a random number in Python using the random() function.

Code:

```
import random
num = random.random()
print(num)
```

Different Ways to Generate a Random Number in Python

There are a number of ways to generate a random numbers in Python using the functions of the Python random module. Let us see a few different ways.

Generating a Random number using choice()

Python random.choice() is an inbuilt function in the Python programming language that returns a random item from a list, tuple, or string.

Code:

```
# import random
import random

# prints a random value from the list
list1 = [1, 2, 3, 4, 5, 6]
print(random.choice(list1))

# prints a random item from the string
string = "striver"
print(random.choice(string))
```

Generating a Random Number using randrange()

The random module offers a function that can generate Python random numbers from a specified range and also allows room for steps to be included, called randrange().

Code:

```
# importing "random" for random operations
import random

# using choice() to generate a random number from a
# given list of numbers.
print("A random number from list is : ", end="")
print(random.choice([1, 4, 8, 10, 3]))

# using randrange() to generate in range from 20
# to 50. The last parameter 3 is step size to skip
# three numbers when selecting.
print("A random number from range is : ", end="")
print(random.randrange(20, 50, 3))
```

Generating a Random number using seed()

Python random.seed() function is used to save the state of a random function so that it can generate some random numbers in Python on multiple executions of the code on the same machine or on different machines (for a specific seed value). The seed value is the previous value number generated by the generator. For the first time when there is no previous value, it uses the current system time.

```
# importing "random" for random operations
import random

# using random() to generate a random number
# between 0 and 1
print("A random number between 0 and 1 is : ", end="")
print(random.random())

# using seed() to seed a random number
random.seed(5)

# printing mapped random number
print("The mapped random number with 5 is : ", end="")
print(random.random())

# using seed() to seed different random number
random.seed(7)

# printing mapped random number
print("The mapped random number with 7 is : ", end="")
print(random.random())

# using seed() to seed to 5 again
```

```
random.seed(5)

# printing mapped random number
print("The mapped random number with 5 is : ", end="")
print(random.random())

# using seed() to seed to 7 again
random.seed(7)

# printing mapped random number
print("The mapped random number with 7 is : ", end="")
print(random.random())
```

Generating a Random number using shuffle()

The shuffle() function is used to shuffle a sequence (list). Shuffling means changing the position of the elements of the sequence. Here, the shuffling operation is in place.

Code:

```
# import the random module
import random


# declare a list
sample_list = ['A', 'B', 'C', 'D', 'E']

print("Original list : ")
print(sample_list)

# first shuffle
random.shuffle(sample_list)
print("\nAfter the first shuffle : ")
print(sample_list)

# second shuffle
random.shuffle(sample_list)
print("\nAfter the second shuffle : ")
print(sample_list)
```

Generating a Random number using uniform()

The uniform() function is used to generate a floating point Python random number between the numbers mentioned in its arguments. It takes two arguments, lower limit(included in generation) and upper limit(not included in generation).

Code:

```
# Python code to demonstrate the working of
```

```
# shuffle() and uniform()

# importing "random" for random operations
import random

# Initializing list
li = [1, 4, 5, 10, 2]

# Printing list before shuffling
print("The list before shuffling is : ", end="")
for i in range(0, len(li)):
    print(li[i], end=" ")
print("\r")

# using shuffle() to shuffle the list
random.shuffle(li)

# Printing list after shuffling
print("The list after shuffling is : ", end="")
for i in range(0, len(li)):
    print(li[i], end=" ")
print("\r")

# using uniform() to generate random floating number in range
# prints number between 5 and 10
print("The random floating point number between 5 and 10 is : ", end="")
print(random.uniform(5, 10))
```

CONCLUSION

What is a Random Variable?

A random variable in statistics are the variables that represent all the possible outcome of a Random Variable.

What are Two Types of Random Variable?

There are two types of Random Variables and that are,

- Continuous Random Variable
- Discrete Random Variable

What is Random Variables Expected Value?

Expected value of a Random Variable is the weighted average of all possible values of the variable. Weight of the random variable is the probability of random variable at specific values.

What are Continuous Random Variables?

Continuous Random Variables are type of random variable in probability theory and statistics that are used to represent the continuous probability of the distribution of a function.