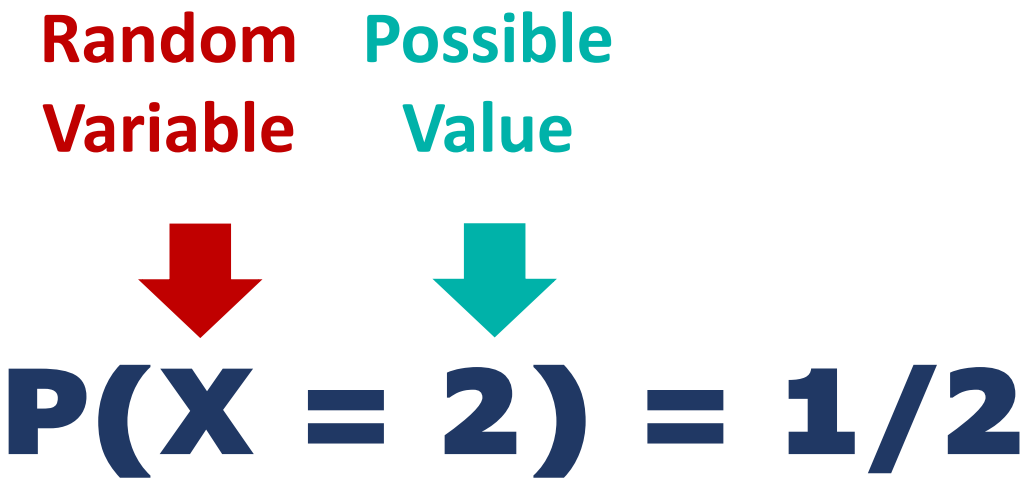
# Random Variables and Probability Distribution Functions

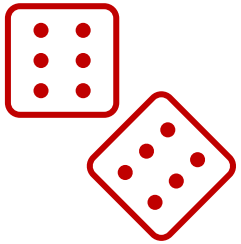
# What is a Random Variable?

A random variable is the set of **possible values** for a random experiment.

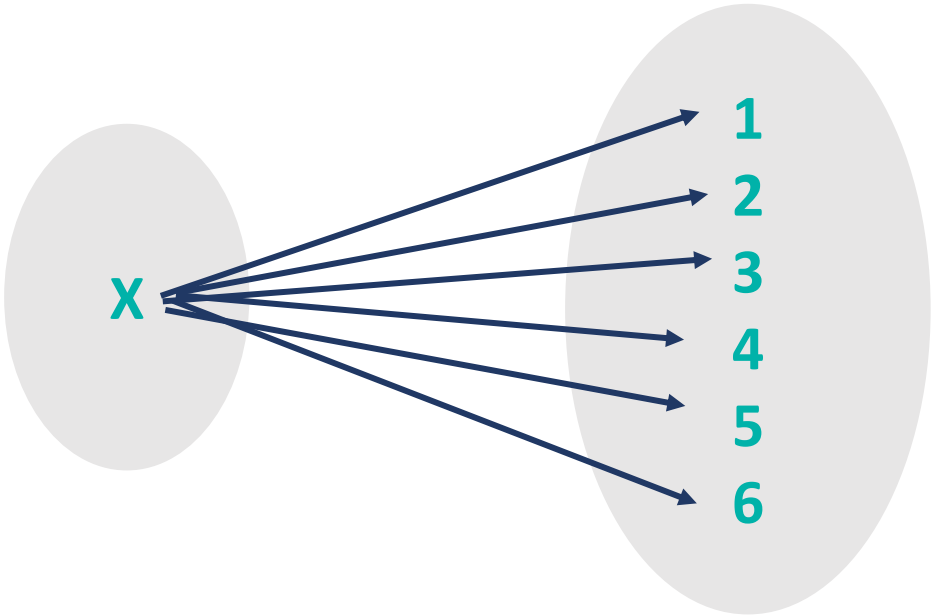



A random variable's set of values is the **sample space.**

# Random Variables Example 1: Rolling a Die

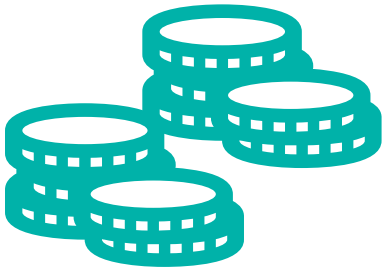


When throwing a six-faced die, we have 6 possible outcomes: 1,2,3,4,5,6. We can define the random variable X associated with this experiment:

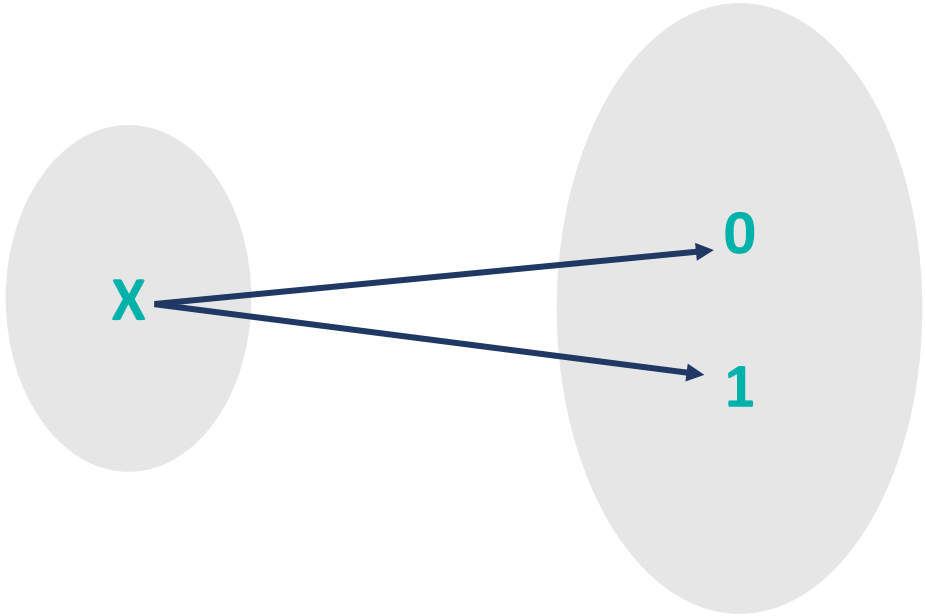



The size of the **random variable is equal to the size of the sample space** for this experiment.

# Random Variables Example: Tossing a Coin



We have 2 possible outcomes when tossing a coin, **Heads=0** and **Tails=1**.
We define the random variable X associated with this experiment below:




Notice that **2, the size of the random variable, is also equal to the size of the sample space** for this experiment

# Computing a Probability Using Random Variables

We can use Random Variables to express the probability of a certain event happening as follows:

**P(X = value) = probability of that value**

## Computing Probability: Rolling Die Example

If we throw a fair die once, the probability of getting any face is 1/6.

| | |
|---|---|
| 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 |

Note that the sum of the probabilities equals 1, as it should be.

# Discrete vs Continuous Random Variables

**Discrete Random Variables**

Random variables that take a **finite** set of values.

**Continuous Random Variables**

Random variables that can take an **infinite** number of possible values.

# Continuous Random Variables Example: Measuring Lake Depth

How deep is a lake within a given radius?

We want to find the range of depth measurements of a one-mile radius around the center of Lake Michigan.

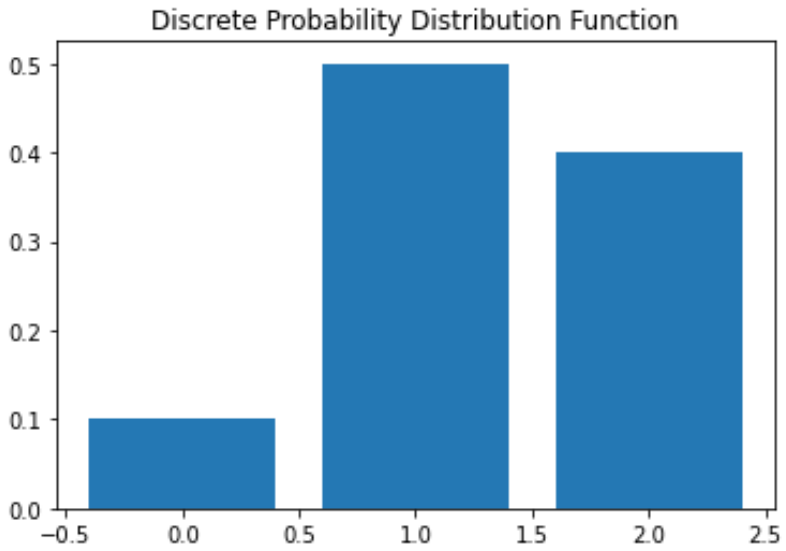
Because the variable depth can take any value, X is a **continuous random variable**.

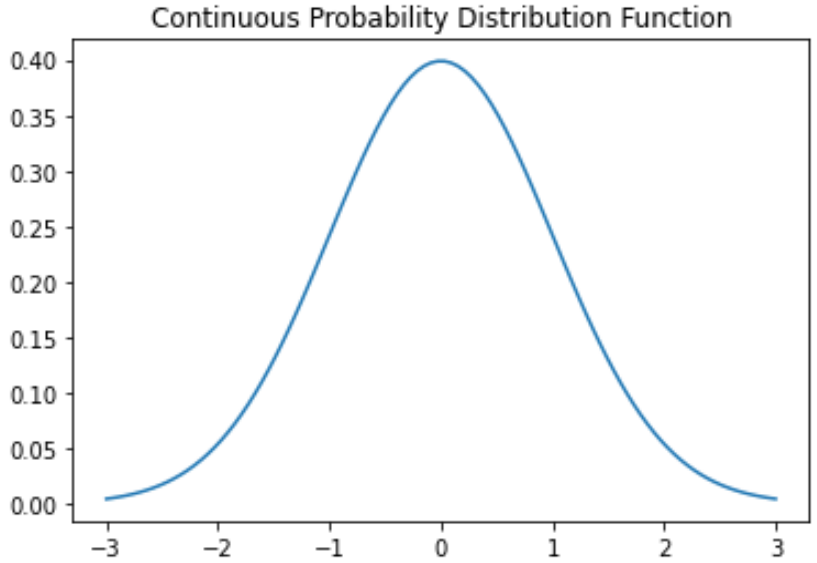
The range for X is the minimum depth possible to the maximum depth possible.

# Probability Distribution Functions

A **probability distribution** is a list of all the possible outcomes of a random variable along with their corresponding probability values.

As with random variables, we can have **discrete and continuous** probability distribution functions.



# Discrete Probability Distribution Functions

Suppose you flip a coin two times.

**X = the number of Heads**

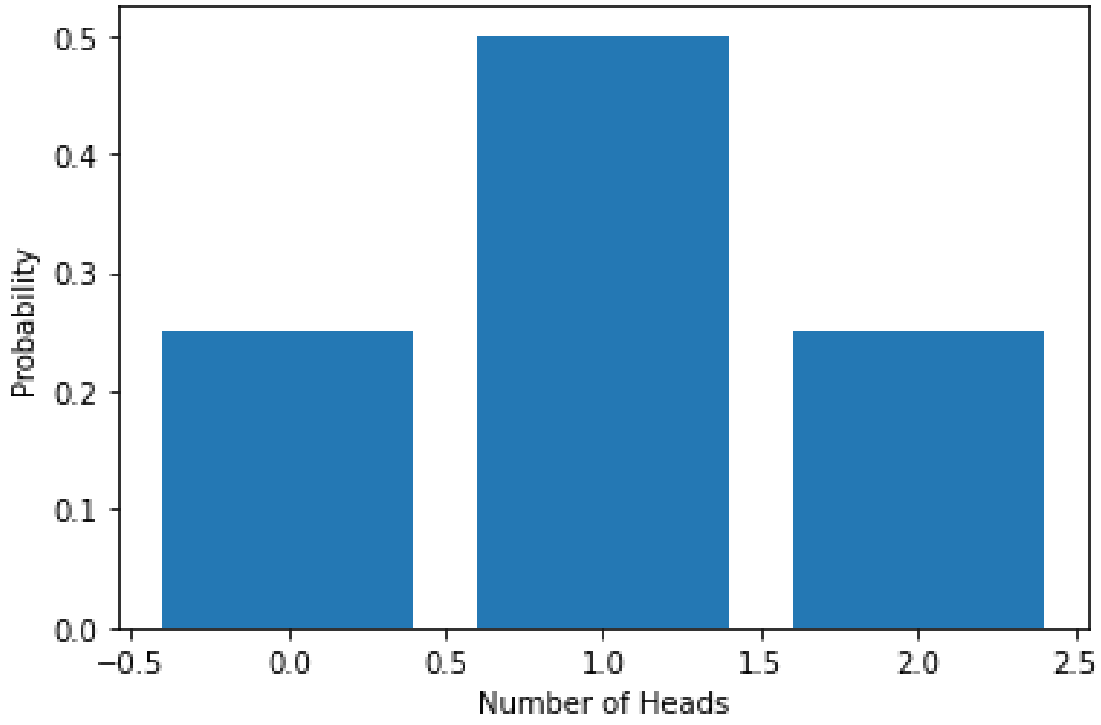
*Hint: The random variable X can only take on the values 0, 1, or 2.*

The probability distribution of this experiment is given by the following table:

| **Outcome (# of Heads) of tossing a coin twice** | 0 | 1 | 2 |
|---|---|---|---|
| **Probability** | 0.25 | 0.5 | 0.25 |

# Discrete Probability Distribution Functions

The probability distribution of tossing a coin twice can also be represented by the following histogram:




- The sum of the area of the single bins in the histogram equals 1, as expected.

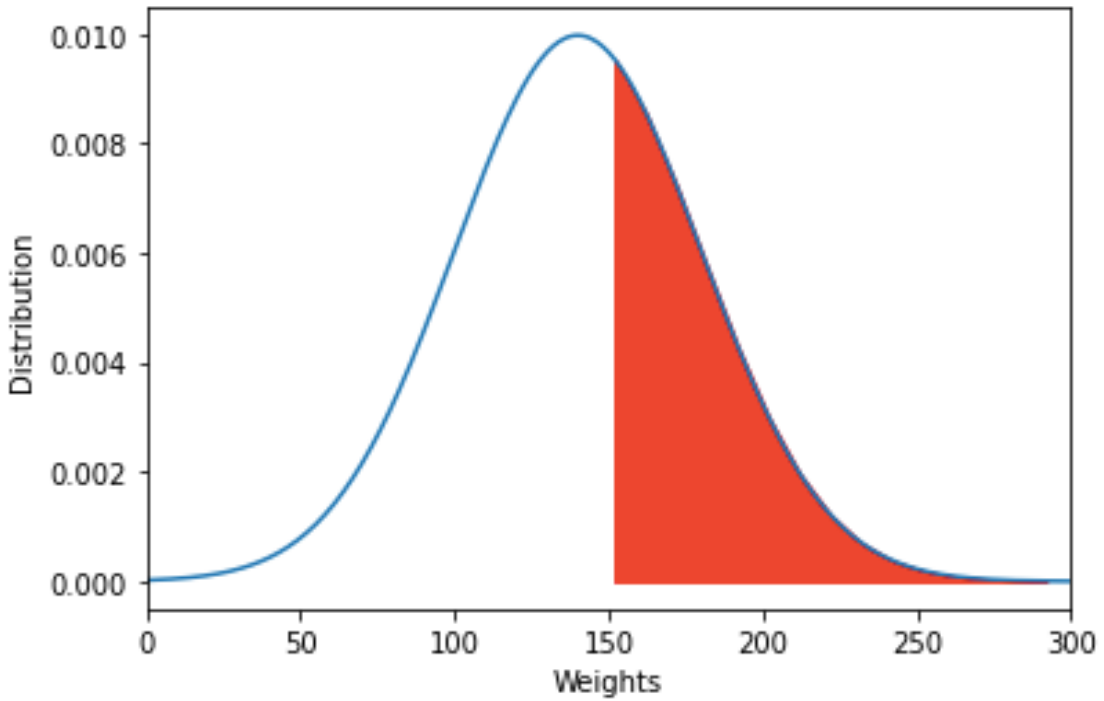
This distribution follows a trend which is also characteristic of a very popular discrete probability distribution function: the **Binomial Distribution**

# Continuous Probability Distribution Functions

- Because they can take **infinitely many** values, the probability that a continuous random variable will assume a particular value is zero. Instead, we compute a probability in terms of **intervals**.
- As a result, a continuous probability distribution cannot be expressed in tabular form.
- Instead, an **equation or formula** is used to describe a continuous probability distribution.

# Continuous Probability Distribution Functions

The curve below represents the probability distribution function of the weight of European women.




The probability that a continuous random variable falls in the interval between *a* and *b* is equal to the area under the curve between *a* and *b*.

For example, the shaded area shows the probability of a weight being > 152 which is 0.38

This curve has bell-shaped trend which is also characteristic of a very popular continuous probability distribution function: the **Normal (or Gaussian) Distribution**

# Summary

- A **random variable** is the set of possible values for a random experiment. Its size, equal the size of the sample space

- We can use random variables to express the probability of a certain event happening

- Random variables can be **discrete or continuous.**

- A **probability distribution** is a list of all the possible outcomes of a random variable along with their corresponding probability values. We can have discrete and continuous probability distribution functions.

•Discrete probability distribution function can be represented in tabular form. Continuous probability distribution function can be expressed as a formula or equation.