



Intro to Probability



Experiments and Events

Learning Objectives

- Be able to recognize what a **sample space** is
- Be able to recognize that an **event** is a subset of the sample space

Data scientists run experiments

- **Experiments**

- are performed following a specific process or set of rules
- can be run many times and have multiple possible outcomes

Data scientists run experiments

- **Experiments**
 - are performed following a specific process or set of rules
 - can be run many times and have multiple possible outcomes
- The **sample space** is the set of all possible outcomes of an experiment

Data scientists run experiments

- **Experiments** are performed following a specific process or set of rules
- Experiments are run many times and have multiple possible outcomes
- The **sample space** is the set of all possible outcomes of an experiment
- An **event** is a set of outcomes of an experiments
 - An event is a subset of the sample space

Data scientists run experiments

- **Experiments** are performed following a specific process or set of rules
- Experiments are run many times and have multiple possible outcomes
- The **sample space** is the set of all possible outcomes of an experiment
- An **event** is a set of outcomes of an experiments
 - An event is a subset of the sample space

Example Experiment: Dice Rolling

Suppose a data scientist wants to record the results of rolling a six-sided die.

The **sample space** is $\{1, 2, 3, 4, 5, 6\}$.

- The **event** for rolling an even number is $\{2, 4, 6\}$
- The **event** for rolling an odd number is $\{1, 3, 5\}$



Sets and Venn Diagrams

Learning Objectives

- Be able to distinguish if sets are **joint** or **disjoint** via Venn diagrams
- Be able to identify the **union** and **intersection** of two sets via Venn diagrams

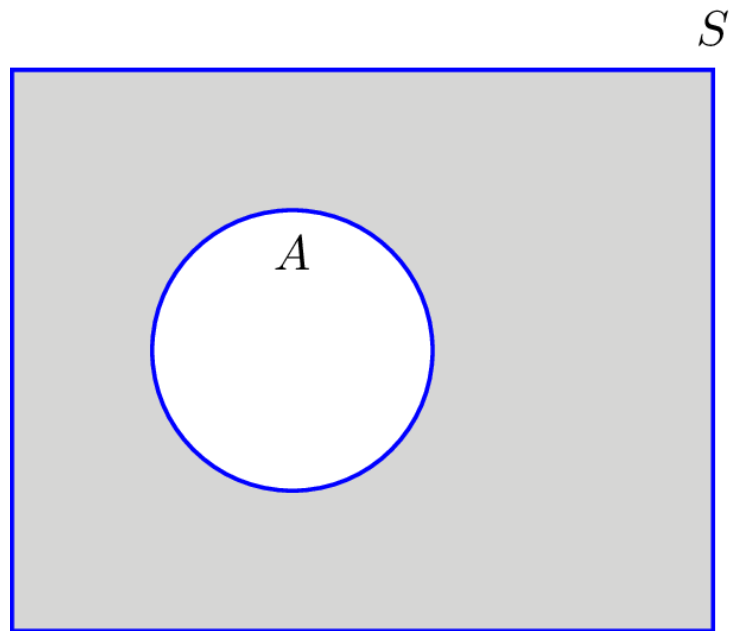
Venn diagrams are used to visualize sets

Denote the **sample space** as S .

- A is an **event** in the sample space
 - A is a subset of S

The sample space not contained within A , i.e. $S - A$, is the **complement** of A .

Example: Let S represent a coin flip. If A represents heads, then the complement of A represents not heads, i.e. tails.



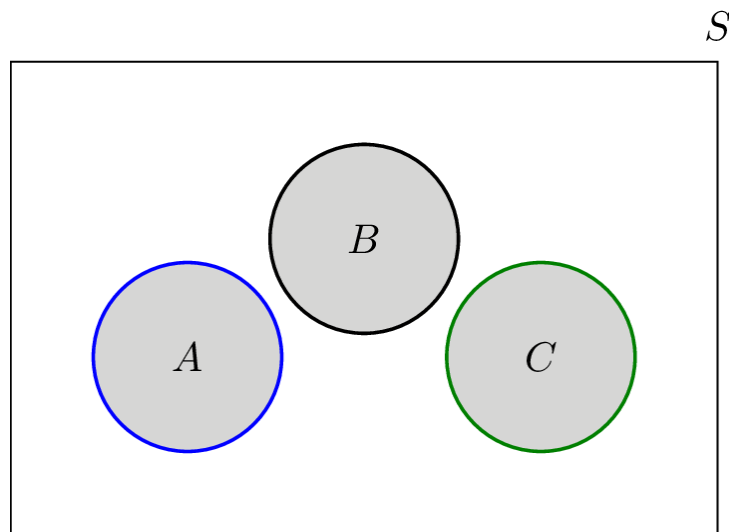
Disjoint sets do not have any shared elements

Let S be the result of a six-sided die roll

- Let A represent the dice rolls $\{1, 2\}$
- Let B represent the dice rolls $\{3, 4\}$
- Let C represent the dice rolls $\{5, 6\}$

A , B , and C are **disjoint**

- Sets that have shared elements (overlap) are called **joint** sets.

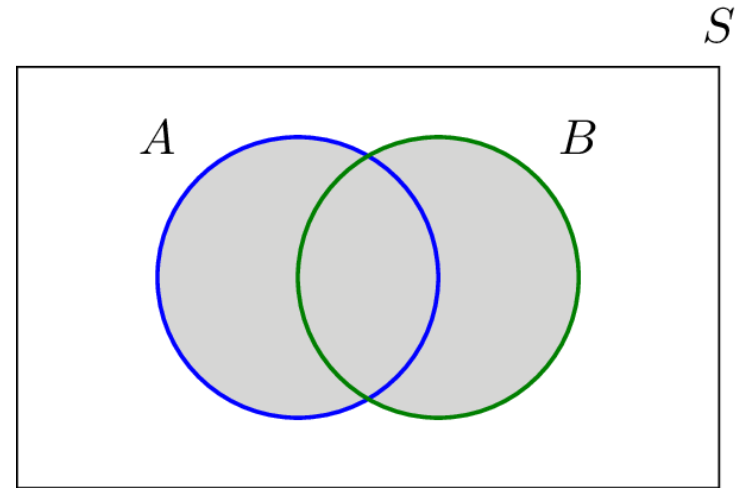


The union of joint sets contains all elements in either set

Let S be the result of a six-sided die roll

- Let A represent an even number roll, i.e. $\{2, 4, 6\}$
- Let B represent a die roll less than 4, i.e. $\{1, 2, 3\}$

Then the $A \cup B$ is $\{1, 2, 3, 4, 6\}$ and the complement of $A \cup B$ is $\{5\}$

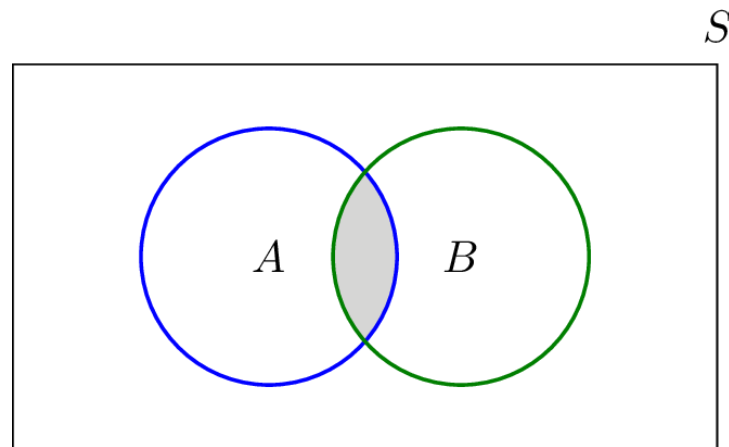


The intersection of joint sets contains all elements in both sets

Let S be the result of a six-sided die roll

- Let A represent an even number roll, i.e. $\{2, 4, 6\}$
- Let B represent a die roll less than 4, i.e. $\{1, 2, 3\}$

Then the intersection of A and B is $\{2\}$

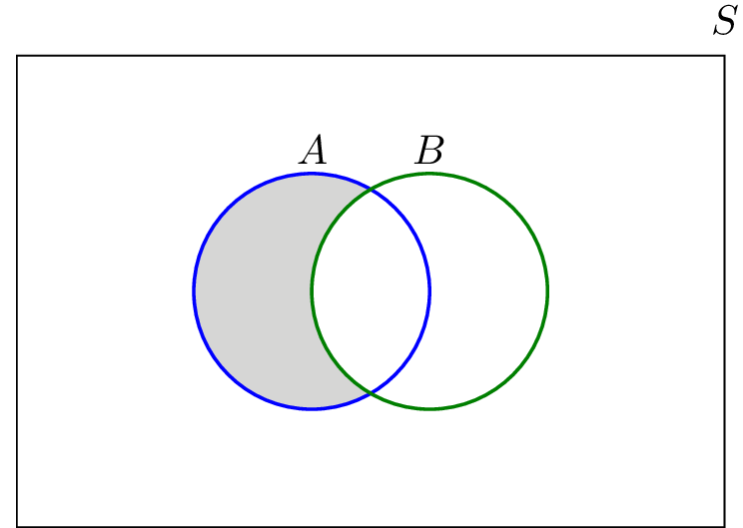


The difference of joint sets, $A - B$, contains the elements of A that are not in B

Let S be the result of a six-sided die roll

- Let A represent an even number roll, i.e. $\{2, 4, 6\}$
- Let B represent a die roll less than 4, i.e. $\{1, 2, 3\}$

Then $A - B$ is $\{4, 6\}$





Probability and Unions of Events

Learning Objectives

- Recognize that probabilities only consists of numbers between 0 and 1
- Be able to calculate probabilities for unions of joint and disjoint events

What is probability?

The probability of an event A happening is

$$P(A) = \text{Number of ways event A can happen} / \text{Total number of possible outcomes}$$

Probabilities fall between 0 and 1.

Probability for Unions of Events

Given two events **A** and **B**, the probability of either happening, i.e. their union is

$$P(\mathbf{A \text{ or } B}) = P(\mathbf{A}) + P(\mathbf{B}) - P(\mathbf{A \text{ and } B})$$

Probability Exercise: Rolling Odd or Even

Given a fair six-sided die:

- Let A represent rolling an odd number, i.e. $\{1, 3, 5\}$
 - Then $P(A) = 3/6 = \frac{1}{2}$
- Let B represent rolling an even number, i.e. $\{2, 4, 6\}$
 - Then $P(B) = 3/6 = \frac{1}{2}$
- Note that A and B are disjoint of each other
 - Then $P(A \text{ and } B) = 0$

Then the probability of rolling an odd or even number is simply

$$P(A \text{ or } B) = P(A) + P(B) - P(A \text{ and } B) = \frac{1}{2} + \frac{1}{2} - 0 = 1$$