The σ**-algebra** $\mathcal{F}$ is the collection of events of interest. Often, $\mathcal{F}$ is the set of all possible events given a sample space Ω, which is just the power set of Ω (Wikipedia 2019g). When referring to an event, we often state that it is an element of $\mathcal{F}$. For instance, we might say an event $A \in \mathcal{F}$.

We're finally ready to assign probabilities to events. We define the **probability measure** $P : \mathcal{F} \to [0, 1]$ to be a function satisfying the following conditions.

1. For every event $A \in \mathcal{F}$, the probability measure of A is greater than or equal to zero—i.e. $P(A) \geq 0$.
2. If an event is the entire sample space, its probability measure is unity—i.e. if $A = \Omega$, $P(A) = 1$.
3. If events $A_1, A_2, \cdots$ are disjoint sets (no elements in common), then $P(A_1 \cup A_2 \cup \cdots) = P(A_1) + P(A_2) + \cdots$.

We conclude the basics by observing four facts that can be proven from the definitions above.

1. $P(\varnothing) = 0$.
2. $P(A \cup B) = P(A) + P(B) - P(A \cap B)$.
3. If $B \subset A$, then $P(B) < P(A)$. In fact, $P(A \setminus B) = P(A) - P(B)$.
4. $P(A_1 \cup A_2 \cup \cdots) \leq P(A_1) + P(A_2) + \cdots$.

3.3 Independence and Conditional Probability

B1

Two events A and B are **independent** if and only if

$$P(A \cap B) = P(A)P(B).$$

If an experimenter must make a judgment without data about the independence of events, they base it on their knowledge of the events, as discussed in the following example.

Example 3.2

Answer the following questions and imperatives.

1. Consider a single fair die rolled twice. What is the probability that both rolls are 6?
2. What changes if the die is biased by a weight such that $P(\{6\}) = 1/7$?
3. What changes if the die is biased by a magnet, rolled on a magnetic dice-rolling tray such that $P(\{6\}) = 1/7$?
4. What changes if there are two dice, biased by weights such that for each $P(\{6\}) = 1/7$, rolled once, both resulting in 6?
5. What changes if there are two dice, biased by magnets, rolled together?

1. Let event $A = \{6\}$. Assuming a fair die, $P(A) = 1/6$. Having no reason to judge otherwise, we assume the results are independent events. Therefore,
$$P(A \cap A) = P(A)P(A) = \frac{1}{6} \cdot \frac{1}{6} = \frac{1}{36}.$$
2. Bias is not dependence. So
$$P(A \cap A) = P(A)P(A) = \frac{1}{7} \cdot \frac{1}{7} = \frac{1}{49}.$$
3. Again, just bias, still independent.
4. Still independent.
5. The magnet dice can influence each other! This means they are not independent! If one wanted to estimate the probability, either a theoretical prediction based on the interaction would need to be developed or several trials could be conducted to obtain an estimation.

3.3.1 Conditional Probability

If events A and B are somehow **dependent**, we need a way to compute the probability of B occurring given that A occurs. This is called the **conditional probability** of B given A, and is denoted $P(B \mid A)$. For $P(A) > 0$, it is defined as

$$P(B \mid A) = \frac{P(A \cap B)}{P(A)}.$$

We can interpret this as a restriction of the sample space Ω to A; i.e. the new sample space $\Omega' = A \subseteq \Omega$. Note that if A and B are independent, we obtain the obvious result:

$$\begin{aligned} P(B \mid A) &= \frac{P(A)P(B)}{P(A)} \\ &= P(B). \end{aligned}$$

Example 3.3

Consider two unbiased dice rolled once. Let events $A = \{\text{sum of faces} = 8\}$ and $B = \{\text{faces are equal}\}$. What is the probability the faces are equal given that their sum is 8?

Directly applying section 3.3.1,

$$
\begin{aligned}
P(B \mid A) &= \frac{P(A \cap B)}{P(A)} \\
&= \frac{P(\{(4,4)\})}{P(\{(4,4)\}) + P(\{(2,6)\}) + P(\{(6,2)\}) + P(\{(3,5)\}) + P(\{(5,3)\})} \\
&= \frac{\frac{1}{6} \cdot \frac{1}{6}}{5 \cdot \frac{1}{6} \cdot \frac{1}{6}} \\
&= \frac{1}{5}.
\end{aligned}
$$

We don't count the event $\{(4,4)\}$ twice, but we do count both $\{(3,5)\}$ and $\{(5,3)\}$, since they are distinct events. We say "order matters" for these types of events.

3.4 Bayes' Theorem

T9

Given two events A and B, **Bayes' theorem** (aka Bayes' rule) states that

$$P(A \mid B) = P(B \mid A)\frac{P(A)}{P(B)}.$$

Sometimes this is written

$$P(A \mid B) = \frac{P(B \mid A)P(A)}{P(B \mid A)P(A) + P(B \mid \neg A)P(\neg A)} \tag{3.1}$$

$$= \frac{1}{1 + \frac{P(B \mid \neg A)}{P(B \mid A)} \cdot \frac{P(\neg A)}{P(A)}}. \tag{3.2}$$

This is a useful theorem for determining a test's effectiveness. If a test is performed to determine whether an event has occurred, we might as questions like "if the test indicates that the event has occurred, what is the probability it has actually occurred?" Bayes' theorem can help compute an answer.

3.4.1 Testing Outcomes

The test can be either positive or negative, meaning it can either indicate or not indicate that A has occurred. Furthermore, this result can be either *true* ☺ or *false* ☹.

There are four options, then. Consider an event A and an event that is a test result B indicating that event A has occurred. table 3.1 shows these four possible test outcomes. The event A occurring can lead to a true positive or a false negative, whereas $\neg A$ can lead to a true negative or a false positive.