3.2 Basic Probability Theory

The mathematical model for a class of measurements is called the **probability space** and is composed of a mathematical triple of a sample space Ω, σ-algebra $\mathcal{F}$, and probability measure P, typically denoted $(\Omega, \mathcal{F}, P)$, each of which we will consider in turn (Wikipedia 2019g).

The **sample space** Ω of an experiment is the set representing all possible **outcomes** of the experiment. If a coin is flipped, the sample space is $\Omega = \{H, T\}$, where H is *heads* and T is *tails*. If a coin is flipped twice, the sample space could be

$$\Omega = \{HH, HT, TH, TT\}.$$

However, *the same experiment can have different sample spaces*. For instance, for two coin flips, we could also choose

$$\Omega = \{\text{the flips are the same, the flips are different}\}.$$

We base our choice of Ω on the problem at hand.

An **event** is a subset of the sample space. That is, an event corresponds to a yes-or-no question about the experiment. For instance, event A (remember: $A \subseteq \Omega$) in the coin flipping experiment (two flips) might be $A = \{HT, TH\}$. A is an event that corresponds to the question, “Is the second flip different than the first?” A is the event for which the answer is “yes.”

3.2.1 Algebra of Events

Because events are sets, we can perform the usual set operations with them.

Example 3.1

Consider a toss of a single die. We choose the sample space to be $\Omega = \{1, 2, 3, 4, 5, 6\}$. Let the following define events.

$$A \equiv \{\text{the result is even}\} = \{2, 4, 6\}$$

$$B \equiv \{\text{the result is greater than 2}\} = \{3, 4, 5, 6\}.$$

Find the following event combinations:

$$A \cup B \quad A \cap B \quad A \setminus B \quad B \setminus A \quad \overline{A} \setminus B.$$

$$A \cup B = \{2, 3, 4, 5, 6\} \quad \text{(even or greater than 2)}$$

$$A \cap B = \{4, 6\} \quad \text{(even and greater than 2)}$$

$$A \setminus B = \{2\} \quad \text{(even but not greater than 2)}$$

$$B \setminus A = \{3, 5\} \quad \text{(greater than two and odd)}$$

$$\overline{A} \setminus B = \{1, 3, 5\} \setminus \{3, 4, 5, 6\} \quad \text{(not even and not greater than 2).}$$

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?