# Probability 1

**Review**

An **experiment** is a procedure that results in an **outcome**. The set of all possible outcomes is called the **sample space**, $S$.

Any subset $E \subseteq S$ of possible outcomes is called an **event**.

A **probability distribution** on a countable sample space $S$ is a function $\mathbb{P} : S \to \mathbb{R}^+$ with $\sum_{s \in S} \mathbb{P}(s) = 1$. Such a distribution is called a **discrete** distribution. This definition can be modified to handle uncountable sample spaces, but we don't worry about this.

The probability distribution that assigns equal probability to each outcome in $S$ is called the **uniform distribution** on $S$. For the uniform distribution, any event $E \subseteq S$ has probability $\mathbb{P}(E) = \frac{|E|}{|S|}$.

The complementary event to an event $E$ is the event $\overline{E} = S \setminus E$. By the properties of a probability distribution, $\mathbb{P}(\overline{E}) = 1 - \mathbb{P}(E)$.

**Conditioning** is a technique for updating the probability of an event given new information by reducing the size of the sample space.

The **conditional probability** of an event $A$ given another event $B$ is given by $\mathbb{P}(A|B) = \frac{\mathbb{P}(A \cap B)}{\mathbb{P}(B)}$.

Two event $A$ and $B$ are **independent** if the occurrence (or not) of one event does not affect the probability of the other. Mathematically, events are independent when $\mathbb{P}(A \cap B) = \mathbb{P}(A)\mathbb{P}(B)$.

A set of events $A_1, A_2, \ldots, A_n$ are **pairwise independent** if each pair of events is independent. A stronger condition is **mutual independence**, where $\mathbb{P}\left(\bigcap_{i \in I} A_i\right) = \prod_{i \in I} \mathbb{P}(A_i)$ for each $I \subseteq \{1, \ldots, n\}$.

Two events $A$ and $B$ are **disjoint** if $A \cap B = \emptyset$.

Suppose that $A_1, \ldots, A_n$ are events which partition the sample space (the events are pairwise disjoint and have union $S$). The **Law of Total Probability** tells us that the probability of another event $B$ is a weighted sum of the conditional probabilities of $B$ given each of the events $A_i$:

$$\mathbb{P}(B) = \mathbb{P}(B|A_1)\mathbb{P}(A_1) + \ldots + \mathbb{P}(B|A_n)\mathbb{P}(A_n).$$

**Bayes' Theorem** is a tool that we can use to reverse the conditioning in a conditional probability. It gives,

$$\mathbb{P}(A_i|B) = \frac{\mathbb{P}(A_i \cap B)}{\mathbb{P}(B)} = \frac{\mathbb{P}(B|A_i)\mathbb{P}(A_i)}{\mathbb{P}(B)} = \frac{\mathbb{P}(B|A_i)\mathbb{P}(A_i)}{\mathbb{P}(B|A_1)\mathbb{P}(A_1) + \ldots + \mathbb{P}(B|A_n)\mathbb{P}(A_n)}.$$

**1.** Consider an experiment where a pair of fair, 6-sided dice, 1 red and 1 green, are rolled.

(a) Describe the sample space of this experiment.

(b) For each of the following events, describe it as a subset of the sample space, and compute its probability.

(i) $E_1 :=$ The red dice shows 3.

(ii) $E_2 :=$ The green dice shows an even number.

(iii) $E_3 :=$ The dice show the same number.

(iv) $E_4 :=$ The sum of the dice rolls is 8.

(v) $E_5 :=$ The numbers on the dice differ by 2.

(vi) $E_6 :=$ The red dice shows a greater number than the green dice.

**2.** Suppose that $A, B$, and $C$ are events (subsets) of a sample space $S$. For each of the following assertions, either argue that it is true using properties of probability, or give an example to show that it is not true.

(a) $\mathbb{P}(A) \leq \mathbb{P}(B) \implies \mathbb{P}(\overline{B}) \leq \mathbb{P}(\overline{A})$.

(b) $\mathbb{P}(A) \leq \mathbb{P}(B) \implies \mathbb{P}(A \cup C) \leq \mathbb{P}(B \cup C)$.

(c) $A \subseteq B \implies \mathbb{P}(A \cup C) \leq \mathbb{P}(B \cup C)$.

(d) $\mathbb{P}(A) \leq \mathbb{P}(B) \implies \mathbb{P}(A \setminus B) \leq \mathbb{P}(B \setminus A)$.

**3.** Four friends, Alice, Bob, Charlie, and Dawn, will pose side-by-side for a picture. Suppose that all possible orderings of the friends are equally likely, and consider the events:

$$E_1 := \text{Charlie is standing next to Dawn.}$$
$$E_2 := \text{Bob is standing next to both Alice and Charlie.}$$
$$E_3 := \text{Alice is standing at one end.}$$

(a) Determine the probabilities of each of these events.

$\mathbb{P}(E_1) =$

$\mathbb{P}(E_2) =$

$\mathbb{P}(E_3) =$

(b) Determine the probabilities of the pairwise intersections of these events.

$\mathbb{P}(E_1 \cap E_2) =$

$\mathbb{P}(E_2 \cap E_3) =$

$\mathbb{P}(E_1 \cap E_3) =$

(c) Compute the following conditional probabilities.

| | | |
|---|---|---|
| $\mathbb{P}(E_1 \mid E_2) :=$ | $\mathbb{P}(E_2 \mid E_3) :=$ | $\mathbb{P}(E_3 \mid E_1) :=$ |
| $\mathbb{P}(E_1 \mid E_3) :=$ | $\mathbb{P}(E_2 \mid E_1) :=$ | $\mathbb{P}(E_3 \mid E_2) :=$ |

(d) Which pairs of events are independent?

**4.** Three identical boxes, each containing 2 coins, are placed on a table. One box contains 2 silver coins, one contains 2 gold coins, and the last contains 1 silver and 1 gold coin. The order of the boxes is chosen uniformly at random.

(a) Suppose that you reach your hand into the leftmost box and randomly pull out a silver coin from it. What is the probability that the other coin in this box is also silver?

(b) **Challenge:** Suppose that after pulling the silver coin out of the leftmost box, you pull a gold coin out of the rightmost box. Now, what is the probability that the other coin in the leftmost box is silver? (**Hint:** What is the sample space of this experiment?)

**5.** A genetic mutation affects 1 in every 1000 individuals (we'll assume that it uniformly and independently affects each individual). A test for this genetic mutation is 99 percent accurate. This means that 99 percent of tests performed on patients with the mutation have a positive result, and 99 percent of tests performed on patients without the mutation have a negative result. Define the events:

$$M := \text{A person has the mutation.}$$
$$T := \text{A person tests positive for the mutation.}$$

(a) Calculate each of the following probabilities and give a brief English description of the event that they describe.

(i) $\mathbb{P}(M) =$

(ii) $\mathbb{P}(\overline{M}) =$

(iii) $\mathbb{P}(T|M) =$

(iv) $\mathbb{P}(T|\overline{M}) =$

(b) Suppose that someone takes the test and it comes back positive. What is the probability that they have the mutation?

(c) Now suppose that they take a second test and it also comes back positive. What is the probability that they have the mutation?