# Discrete Probability and Counting

A *finite probability space* is a set $S$ and a function $p : S \to R_{\geq 0}$ s.t.:

- $p(s) > 0\, \forall s \in S$ and
- $\sum_{s \in S} p(s) = 1$.

We refer to $S$ as the *sample space*, subsets of $S$ as *events* and $p$ as the *probability distribution*. The probability of an event $A \subseteq S$ is $\sum_{a \in A} p(a)$. ($p(\emptyset) = 0$.)

**Example**: Suppose we flip a fair coin. Saying the coin is fair implies that it is equally likely to flip $H$ (heads) or $T$ (tails) therefore $p(H) = p(T) = 1/2$.

If we assign all elements of $S$ the same probability, as in the example above, then $p$ is the *uniform distribution*.

**Example**: Suppose we flip a biased coin where the probability of $H$ is twice as much as the probability of $T$. Since $p(H) + p(T) = 1$, this implies $p(H) = 2/3$ and $p(T) = 1/3$.

**Example:** Suppose we flip a fair coin twice. What is the probability of getting one $H$ and one $T$? All the possible outcomes are $\{HH, HT, TH, TT\}$. Two out of the possible 4 outcomes give us one $H$ and one $T$, each outcome has probability 1/4 therefore the total probability is 1/2

Suppose we flipped a fair coin $n$ times. How many possible outcomes are there? There are two choices for each flip of the coin, so there are $2^n$ possible outcomes. The probability of getting any one of these is $1/2^n$. (Where did we see this before?)

Now suppose we want to know the probability of getting exactly $k$ $H$s. We need to know how many of the $2^n$ strings have exactly $k$ $H$s. In general, the number of ways to choose $k$ things from $n$ is given by:

$\binom{n}{k} = n!/(n-k)!k!$

Where $n! = n(n-1)(n-2)\ldots 1$ and is read $n$ factorial. We define $0! = 1$. Note that $\binom{n}{k} = \binom{n}{n-k}$. These numbers are known as the *binomial coefficients*. Consider $(x+y)^2 = x^2 + 2xy + y^2$. The coefficients of this polynomial are $\{1, 2, 1\}$ which are the numbers $\binom{2}{0}$, $\binom{2}{1}$, $\binom{2}{2}$. In general, $(x+y)^n = \binom{n}{0}x^n + \binom{n}{1}x^{n-1}y + \binom{n}{2}x^{n-2}y^2 + \ldots + \binom{n}{n-1}xy^{n-1} + \binom{n}{n}y^n$.

**Example**: Suppose we flip a fair coin 10 times. What is the probability of getting exactly 4 $H$s? First we compute $\binom{10}{4} = 210$. Then we compute the total number of outcomes $2^{10} = 1024$. Therefore the probability of getting exactly 4 $H$s is $210/1024 \approx .205$

Two events are *disjoint* if their intersection is empty.

**Example**: In the example of flipping 2 coins, the event $A =$ 'getting exactly one $H$' and the event $B =$ 'getting exactly $2H$s' are disjoint. But, $A$ is not disjoint from the event $C =$ 'getting exactly one $T$'. In fact, events $A$ and $C$ are the same in this case.

In general we have: $p(A \cup B) + p(A \cap B) = p(A) + p(B)$. Therefore, for disjoint events we have: $p(A \cup B) = p(A) + p(B)$. The first statement follows from the principle of *inclusion - exclusion* which states that $|A \cup B| = |A| + |B| - |A \cap B|$.

**Example**: Say we flip a coin 10 times. What is the probability that the first flip is a $T$ or the last flip is a $T$? The number of outcomes with the first flip $T$ is $2^9$. The number of outcomes where the last flip is a $T$ is $2^9$. The number of strings with both properties is $2^8$. Hence, the number of strings with either property is $2^9 + 2^9 - 2^8 = 768$.

Suppose we know that one event has happened and then want to ask about another. For two events $A$ and $B$, the *conditional probability* of $A$ relative to $B$ is $p(A|B) = p(A \cap B)/p(B)$ and read the probability of $A$ given $B$.

**Example**: Suppose we flip a fair coin 3 times. Let $B$ be the event that we have at least one $H$ and $A$ be the event of getting exactly 2 $H$s. What is the probability of $A$ given $B$? In this case, $(A \cap B) = A$, $p(A) = 3/8$ (why?), $p(B) = 7/8$ (why?), and therefore $p(A|B) = 3/7$.

Notice that the definition of conditional probability also gives us the formula: $p(A \cap B) = p(A|B)p(B)$. For three events we have: $p(A \cap B \cap C) = p(A|B \cap C)p(B|C)p(C)$. (What is a general rule?)

We can also use conditional probabilities to find the probability of an event by breaking the sample space into disjoint pieces. If $S = S_1 \cup S_2 \ldots \cup S_n$

and all pairs $S_i$, $S_j$ are disjoint then for any event $A$, $p(A) = \sum p(A|S_i)p(S_i)$.

**Example**: Suppose we flip a fair coin twice. Let $S_1$ be the outcomes where the first flip is $H$ and $S_2$ be the outcomes where the first flip is $T$. What is the probability of $A =$ getting 2 $H$s? $p(A) = (1/2)(1/2) + (0)(1/2) = 1/4$.

Two events $A$ and $B$ are *independent* if $p(A \cap B) = p(A)p(B)$. This immediately gives: $A$ and $B$ are independent iff $p(A|B) = p(A)$.

If $p(A \cap B) > p(A)p(B)$ then $A$ and $B$ are *positively correlated*.

If $p(A \cap B) < p(A)p(B)$ then $A$ and $B$ are *negatively correlated*.

**Example**: In the example of flipping 3 coins, $p(A|B) \neq p(A)$ therefore these two events are not independent. Let $C$ be the event that we get at least one $H$ and at least one $T$. Let $D$ be the event that we get at most one $H$. $p(C) = 6/8$, $p(D) = 4/8$, and $p(C \cap D) = 3/8$ therefore events $C$ and $D$ are independent.

We say events $A_1, \ldots A_n$ are *mutually independent* if for all subsets $S \subseteq \{1, \ldots, n\}$, $p(\cap_{i \in S} A_i) = \prod p(A_i)$. (What is an example of a set of mutually independent events?)