5. Discrete Probability

In probability theory, the goal is to determine how likely something is to happen, when studying phenomena that are unpredictable. An *experiment* here means any controlled, physical action that can be repeated an arbitrary number of times. Experiments for which there is more than one possible result are said to be *probabilistic*—or simply *random*. The degree of certainty associated with the result of a random experiment is then what we refer to as "probability."

Sample spaces

A distinct observation that results from a random experiment is called an *outcome*, and a *sample space* is the collection of all possible outcomes of such an experiment. (The sample space is sometimes also called a *universe*.) When an experiment is executed once and an outcome is observed, this is known as a *trial*.

Examples:

Many examples of probabilistic experiments are found in games of chance. For a roll of a single, six-sided die, the sample space may be defined as the set

$$\Omega = \{1, 2, 3, 4, 5, 6\}.$$

Similarly, a coin flip has only two possible outcomes, and is characterized by a sample space

$$\{\text{Head}, \text{Tails}\}, \text{ which can also be encoded as the set } \{0, 1\}.$$

Flipping a coin n times creates a sample space $\{0, 1\}^n$, which is the set of all binary n-strings. A sample space can, unlike these examples, be infinite. Discrete probability theory, however, only deals with finite and countably infinite sample spaces.

Probability distributions

Each outcome in a sample space is associated with a probability, or *weight*, denoted by a real number in the range 0 to 1 (both inclusive), which represents its likelihood to occur. A *probability distribution* (or probability measure) on a sample space Ω is a function $P : \Omega \to [0, 1]$. An important requirement here is that the sum of all probabilities in the sample space must be equal to 1.

$$\text{If } \Omega \text{ is a sample space, then } \sum_{i \in \Omega} P(i) = 1.$$

We can also think of the probability of an outcome as the proportion of times that it occurs in a random experiment, when the experiment is repeated a sufficiently large number of times. If $C_i(n)$ denotes the the number of occurences of i observed when sampling at random n times from a probability distribution P, then

$$P(i) = \lim_{n \to \infty} \frac{C_i(n)}{n}.$$

The uniform distribution

If we assume a fair die in our first example, we obtain a distribution which is said to be *uniform*:

i	1	2	3	4	5	6
$P(i)$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$

A uniform probability distribution is one that satisfies the property

$$\forall i \in \Omega \left[P(i) = \frac{1}{|\Omega|} \right].$$

Bernoulli distribution

TODO

Point distribution

TODO

Events

An *event* is simply a subset of a sample space. If S is an event defined on the sample space Ω, then

$$S \subseteq \Omega.$$

For some concrete examples, we can think of a lottery with exactly one winner, and the following set of participants (i.e., the sample space)

$$\{\text{Angela, Donald, Justin, Kim, Robert, Stefan, Theresa}\}.$$

Then, the event

1. of the winner being male is {Donald, Justin, Kim, Robert, Stefan},
2. that the winner's name contains the letter 'E' is {Angela, Robert, Stefan, Theresa},
3. that the winner is Justin is {Justin}, and
4. that the winner is **not** Justin is {Angela, Donald, Kim, Robert, Stefan, Theresa}.

(TO BE CONTINUED)