Probability Review

We begin the discussion by reviewing some definitions and concepts from probability theory. These definitions and concepts are required to ensure we are using a common vocabulary and have the basic set of tools necessary to discuss algorithms probabilistically. We begin by considering a set of items $\mathbf{S}$.

Definition: A *sample space* is a set $\mathbf{S}$ whose elements are called elementary events.

Definition: An *elementary event* can be treated as a possible outcome of an experiment.

Definition: An *event* is a subset of $\mathbf{S}$.

Definition: Two events, $\mathbf{A}$ and $\mathbf{B}$, are *mutually exclusive* if and only if $\mathbf{A} \cap \mathbf{B} = \emptyset$.

By definition, all elementary events are mutually exclusive.

Probability Distributions

In the following, we will use $P(a)$ to denote the probability of event a. Then we say that a *probability distribution* $P()$ on $\mathbf{S}$ is a mapping from events in $\mathbf{S}$ to real numbers $\Re$ such that the following holds:

- $P(a) \geq 0$ for all $a \subseteq \mathbf{S}$,
- $P(\mathbf{S}) = 1$, and
- $P(a \cup b) = P(a) + P(b)$ for any two mutually exclusive events a and b.

Note that if the events a and b are not mutually exclusive, then we have $P(a \bigcup b) = P(a) + P(b) - P(a \bigcap b)$.

Given the axioms of probability and the notion of a probability distribution, we are now prepared to consider several types and examples of probability distributions. Generally, we are concerned with two types of distributions—discrete and continuous. A probability distribution is said to be *discrete* if it is defined over a finite (or countably infinite) sample space. A probability distribution is said to be *continuous*, on the other hand, if the sample space is neither finite nor countably infinite. When considering these distributions, we refer to the associated probability functions as probability *mass* functions and probability *density* functions respectively.

For a particular discrete probability distribution over a sample space $\mathbf{S}$ with event a, we can define the following:

$$P(a) = \sum_{s \in a \subseteq \mathbf{S}} P(s).$$

Thus, the probability distribution assigns a probability value to every event (elementary or otherwise) in the sample space $\mathbf{S}$. For example, suppose $\mathbf{S}$ is finite and every has a probability $P(s) = 1/|\mathbf{S}|$, where s is an elementary event. This particular distribution is called a *uniform* distribution. Two common examples of sample spaces with uniform probability distributions are coin flips and die tosses. If one has a fair coin with two sides, then $\mathbf{S} = \{\text{Heads, Tails}\}$, and $P(\text{Heads}) = P(\text{Tails}) = 1/2$. If one has a fair die with six sides, then $\mathbf{S} = \{1, 2, 3, 4, 5, 6\}$, and $P(1) = P(2) = P(3) = P(4) = P(5) = P(6) = 1/6$.

For continuous probability distributions, it is common that these distributions are defined over ranges of values (typically closed intervals0 since the probability of any specific value is zero. Consider the uniform probability distribution for a continuous sample space. Then consider two closed intervals, $[a, b] \in \mathbf{S}$ and $[c, d] \in \mathbf{S}$, where $[c, d] \in [a, b]$, then $P([c, d]) = (d - c)/(b - a)$. An interesting corollary to working with intervals over a uniform distribution is that we can define events to be any subset of $[a, b]$ obtained by a finite or countable union of either open or closed intervals. Note specifically that for any $a, b, \in \mathbf{S}$, $P([a, a]) = P([b, b]) = 0$. Note also that the definition of a closed interval permits use to rewrite that interval as $[a, b] = [a, a] \cup (a, b) \cup [b, b]$. Therefore, we can substitute in the point probabilities to show that $P([a, b]) = P((a, b))$.