Randomized Algorithms and Probabilistic Techniques

September 15, 2021

1 08/24

This course will involve only a few techniques with a variety of applications. Many domains, such as Distributed Computing, Networks, etc., all depend on randomized algorithms.

We will focus on a few techniques that lead to understanding.

1. Union Bound
2. Linearity of Expectation
3. Markov's Inequality
4. Chernoff Bounds

These four things prove very useful in the design and understanding of algorithms.

1.1 Basic Definitions

Definition 1. *A* Sample Space *is a set S whose elements consist of* simple events *(also called* elementary *events). When S is finite or countably infinite, we say it is a* Discrete *sample space.*

Definition 2. *An* event *in a sample space S is a subset of S.*

Definition 3. *A* Probability Distribution *on S is a function $\mathbb{P} : 2^S \to [0, 1]$ that satisfies*

1. $\mathbb{P}(S) = 1$
2. *If E_1, E_2, ... are pairwise disjoint events (i.e., $E_i \cap E_j = \emptyset$ for all pairs i, j, also called mutually exclusive), indexed by some finite or countably infinite set I, then*

$$\mathbb{P}\left(\bigcup_{i \in I} E_i\right) = \sum_{i \in I} \mathbb{P}(E_i)$$

1.2 Conditional Probability

The expression $\mathbb{P}(E_2 \mid E_1)$ is read "the probability of E_2 given E_1." For example, if we randomly select a person from Texas, we might write

$$\begin{aligned} E_1 &= \text{person chosen is in Houston} \\ E_2 &= \text{person chosen is a UH student} \end{aligned}$$

in which case $\mathbb{P}(E_2 \mid E_1)$ is simply the probability that a randomly selected person from Texas is a UH student *given that they are in Houston.*

In general, the conditional probability $\mathbb{P}(E_2 \mid E_1)$ is defined

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

which intuitively can be thought of as taking the probability that E_2 and E_1 occur and "normalizing it" by dividing by the probability that E_1 occurs.

1.3 Independence

Two events, E_1 and E_2, are *independent* if

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

or, equivalently, if

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

For example, suppose C_1 and C_2 are the outcomes of two fair coin tosses. These are independent, since

$$\mathbb{P}(C_1 = H \cap C_2 = T) = \mathbb{P}(C_1 = H)\,\mathbb{P}(C_2 = T) = \frac{1}{4}$$

Independence is not always related to physical independence. For example, say we are given a fair die and let

$$\begin{aligned}
E_1 &= \text{roll is even}\\
E_2 &= \text{roll is less than or equal to 4}
\end{aligned}$$

In this case, we can enumerate the sample space and explicitly determine $\mathbb{P}(E_1)$, $\mathbb{P}(E_2)$, $\mathbb{P}(E_1 \cap E_2)$, and $\mathbb{P}(E_1)\,\mathbb{P}(E_2)$, to see if the events are independent:

$$\begin{aligned}
E_1 &= \{2,4,6\}\\
E_2 &= \{1,2,3,4\}\\
E_1 \cap E_2 &= \{2,4\}
\end{aligned}$$

Then

$$\begin{aligned}
\mathbb{P}(E_1) &= \frac{3}{6} = \frac{1}{2}\\
\mathbb{P}(E_2) &= \frac{4}{6} = \frac{2}{3}\\
\mathbb{P}(E_1 \cap E_2) &= \frac{2}{6} = \frac{1}{3}\\
\mathbb{P}(E_1)\,\mathbb{P}(E_2) &= \frac{1}{2} \cdot \frac{2}{3} = \frac{1}{3}
\end{aligned}$$

Thus, we see the events are independent. On the other hand, if E_2 is the event that the roll is *strictly less* than 4, we have

$$\begin{aligned}
E_1 &= \{2,4,6\}\\
E_2 &= \{1,2,3\}\\
E_1 \cap E_2 &= \{2\}
\end{aligned}$$

Then

$$\begin{aligned}
\mathbb{P}(E_1) &= \frac{3}{6} = \frac{1}{2}\\
\mathbb{P}(E_2) &= \frac{3}{6} = \frac{1}{2}\\
\mathbb{P}(E_1 \cap E_2) &= \frac{1}{6}\\
\mathbb{P}(E_1)\,\mathbb{P}(E_2) &= \frac{1}{2} \cdot \frac{1}{2} = \frac{1}{4}
\end{aligned}$$

and we see that the events are *not independent.*

1.4 The Inclusion-Exclusion Principle

A basic result in set theory is that

$$\begin{aligned}
|A \cup B| &= |A| + |B| - |A \cap B|\\
|A \cup B \cup C| &= |A| + |B| + |C| - |A \cap B| - |A \cap C| - |B \cap C| + |A \cap B \cap C|
\end{aligned}$$

which can be generalized to an arbitrary finite union by

$$\bigcup_{i=1}^{n} A_i = \sum_{k=1}^{n} (-1)^{k+1} \left(\sum_{1 \le i_1 < \cdots < i_k \le n} |A_{i_1} \cap \cdots \cap A_{i_k}| \right)$$

or equivalently

$$\left| \bigcup_{i=1}^{n} A_i \right| = \sum_{\emptyset \neq J \subseteq \{1,\ldots,n\}} (-1)^{|J|+1} \left| \bigcap_{j \in J} A_j \right|$$

This yields the probability formulas

$$\mathbb{P}(E_1 \cup E_2) = \mathbb{P}(E_1) + \mathbb{P}(E_2) - \mathbb{P}(E_1 \cap E_2)$$
$$\mathbb{P}\left(\bigcup_{i=1}^{n} E_i\right) = \sum_{\emptyset \neq J \subseteq \{1,\ldots,n\}} (-1)^{|J|+1}\, \mathbb{P}\left(\bigcap_{j \in J} E_j\right)$$

1.5 Union Bound

In general
$$\mathbb{P}\left(\bigcup E_i\right) \leq \sum \mathbb{P}(E_i)$$
While this bound is often not very precise, it is useful in many cases where the events E_i are "bad" and we can bound the likelihood of a single E_i. This allows us to bound the likelihood of *any* E_i.

1.6 Conditioning on Multiple Events (Chain Rule)