Discussion 7B Recap

Tyler Zhu

April 4, 2020

1 Introduction to Discrete Probability

We introduce the key foundations for talking about probability, so this discussion is mostly about becoming familiar with the concept.

Definition 1 (Probability Space). A **probability space** is a *sample space* Ω along with a probability function $\mathbb{P}(\omega)$ acting on each sample point ω which satisfies two properties (the Kolmogorov axioms):

- (Non-negativity) $\mathbb{P}(\omega) \geq 0$ for all sample points $\omega \in \Omega$, and
- (Totality) $\sum_{\omega} \mathbb{P}(\omega) = 1$, i.e. the sum of all of the probabilities is 1.

Definition 2 (Event). An **event** in a probability space $(\Omega, \mathbb{P})$ is a subset $A \subseteq \Omega$ of the sample space. The probability of such an event is the sum of the probabilities of all the samples in A, i.e. $\mathbb{P}(A) = \sum_{\omega \in A} \mathbb{P}(\omega)$.

This is a more formal treatment than you may have come across in other contexts. Usually we talk about probabilities without reference to the sample space, but for our future discussion this will be important.

The important idea to keep around is that we have a collection of sample points (I like to call them outcomes) which are indivisible and together form our sample space. We can then assign probabilities of each outcome happening (which themselves must be in between 0 and 1), and that forms a probability space. So really a probability space is just a pair $(\Omega, \mathbb{P})$.

This is useful because I can easily define two different probability functions on the same sample space. If $\Omega_1 = \{H, T\}$ is the set of all outcomes of flipping a coin, then I can model having a fair coin by picking a function $\mathbb{P}_1$ for which $\mathbb{P}_1(H) = \mathbb{P}_1(T) = \frac{1}{4}$. We would call such a space *uniform*, since all outcomes have the same probability. But I can also model having a biased coin by picking a function $\mathbb{P}_2$ for which $\mathbb{P}_2(H) = \frac{3}{4}, \mathbb{P}_2(T) = \frac{1}{4}$.

Finally, we call sets of sample points **events**, and these are just the different scenarios that one might be interested in under our probability space. For example, if I have $\Omega_2 = \{HH, HT, TH, TT\}$ be the set of all outcomes of flipping two coins, some events could be $A =$ flipping at least one Head, $B =$ flipping two tails, and so on. In this case, $A = \{HH, HT, TH\}$ and $B = \{TT\}$. We can define probabilities on events by simply summing the probabilities of all of its sample points. So $\mathbb{P}(A) = \frac{3}{4}$ and $\mathbb{P}(B) = \frac{1}{4}$.