4 Probability Distributions

Throughout the past chapters, we've actually already encountered many of the topics in this section. In order to define things like expectation and variance, we introduced random variables denoted X or Y as mappings from the sample space to the real numbers. All of the distributions we've so far looked at have been what are called *discrete* distributions. We will soon look at the distinction between discrete and continuous distributions. Additionally we will introduce perhaps the most influential theorem in statistics, the *Central Limit Theorem*, and give some applications.

4.1 Random Variables

In Section 2.2 (Expectation), we wanted to find the expectation of a coin flip. Since the expectation is defined as a weighted sum of outcomes, we needed to turn the outcomes into numbers before taking the weighted average. Otherwise, we would end up with weighted averages that look like $\frac{1}{2}\mathtt{H} + \frac{1}{2}\mathtt{T}$, which makes no sense. We provided the mapping

$$\mathtt{T} \mapsto 0$$
$$\mathtt{H} \mapsto 1$$

Here was our first encounter of a random variable.

Definition 4.1. *A function X that maps outcomes in our sample space to real numbers, written $X : \Omega \to \mathbb{R}$, is called a* ***random variable****.*

In the above example, our sample space was

$$\Omega = \{\mathtt{H}, \mathtt{T}\}$$

and our random variable $X : \Omega \to \mathbb{R}$, i.e. our function from the sample space Ω to the real numbers $\mathbb{R}$, was defined by

$$X(\mathtt{T}) = 0$$
$$X(\mathtt{H}) = 1$$

Now would be a great time to go onto the website and play with the "Random Variable" visualization. The sample space is represented by a hexagonal grid. Highlight some hexagons and specify the value your random variable X assigns to those hexagons. Start sampling on the grid to see the empirical frequencies on the left.

4.1.1 Independence of Random Variables

In previous sections we've mentioned independence of random variables, but we've always swept it under the rug during proofs since we hadn't yet formally defined the concept of a random variable. Now that we've done so, we can finally define a second form of independence: the independence of *random variables* (more general than the independence of *events*).

Definition 4.2. *Suppose X and Y are two random variables defined on some sample space Ω. We say X and Y are* ***independent random variables*** *if*

$$P(X \in A, Y \in B) = P(X \in A)P(Y \in B)$$

for any two subsets A and B of $\mathbb{R}$.

Let's go back and prove Exercise 2.9 (c), i.e. that if X and Y are independent random variables, then

$$E[XY] = E[X]E[Y]$$

Proof. Define the random variable $Z(\omega) = X(\omega)Y(\omega)$. By the definition of expectation, the left hand side can be written

$$\begin{aligned}
E[XY] &= \sum_{z \in Z(\Omega)} z \cdot P(Z = z) \\
&= \sum_{x \in X(\Omega), y \in Y(\Omega)} xyP(X = x, Y = y) \\
&= \sum_{x \in X(\Omega)} \sum_{y \in Y(\Omega)} xyP(X \in \{x\}, Y \in \{y\}) \\
&= \sum_{x \in X(\Omega)} \sum_{y \in Y(\Omega)} xyP(X \in \{x\})P(Y \in \{y\}) \\
&= \sum_{x \in X(\Omega)} xP(X \in \{x\}) \sum_{y \in Y(\Omega)} yP(Y \in \{y\}) \\
&= E[X]E[Y]
\end{aligned}$$

This completes the proof. □