Math/Stat 251 Fall 2015
Summary of Lecture from September 16, 2015

"Why do we flip coins so often in this class? Why do we roll so many dice? I realize that these are chance experiments, but surely there are more exciting examples you could come up with rather than dice or coins!"

This is a comment that I often receive from students and I would like to explain why dice and coins are frequently used. Consider the following scenarios.

- A wildlife biologist examines certain fish for a genetic trait he suspects may be linked to sensitivity to industrial toxins in the environment. What is the probability that a randomly selected fish displays sensitivity?

- A computer scientist designing a typical I/O (input/output) subsystem will ensure that there are multiple communication lines open in case a component fails. The CPU (central processing unit) might be attached to two channels. Information is routed through the first channel if that channel is available; otherwise it is routed through the second channel. What is the probability that a randomly selected data stream follows the first channel?

- A simple (but risky) strategy for trading stocks is to buy a stock, hold it for a short period of time, and then sell it. The investor obviously makes money if and only if the stock goes up. Therefore, the investor is interested in the probability that the stock goes up.

- Government political strategists are interested in gauging public support for proposed legislation and the effect that its implementation will have on their party's standing. Eligible voters are contacted and asked whether they support or oppose the legislation. The strategists are interested in the probability that such a voter supports the legislation.

In all these examples, there are two possible outcomes. These can be called *success and failure* or *yes and no* or *0 and 1* or *on and off* or *heads and tails* or etc. For instance,

- fish displays sensitivity vs. fish does not display sensitivity,

- data stream follows first channel vs. data stream follows second channel,

- stock goes up vs. stock goes down,

- voter opposes the legislation vs. voter supports the legislation.

Since there are only two possible outcomes, if a success occurs with probability p, then a failure must occur with probability $1 - p$. It is perhaps worth noting that we use the phrase *Bernoulli trial* to indicate a chance experiment where one of only two distinct outcomes may result.

Hence, if we want to construct a mathematical model of a Bernoulli trial, we can simplify things and consider flipping a coin with probability p of heads and probability $1-p$ for tails. This is why we flip coins so often.

More formally, let $0 \le p \le 1$ be fixed, consider the sample space $S = \{H, T\}$, and let the probability $\mathbf{P}$ be defined by

$$\mathbf{P}(\{H\}) = p \quad \text{and} \quad \mathbf{P}(\{T\}) = 1 - p.$$

Define the random variable X by setting

$$X(H) = 1 \quad \text{and} \quad X(T) = 0$$

so that X counts the number of successes on one trial. Moreover, it follows that

$$\mathbf{P}(X = 1) = p \quad \text{and} \quad \mathbf{P}(X = 0) = 1 - p.$$

We often say that X is a *Bernoulli random variable with parameter* p and write $X \sim$ Bernoulli(p). Note that it is traditional for the values of a Bernoulli(p) random variable to be 0 and 1.

In the case where there are more than two possible outcomes, we might use a die instead! For instance, consider a public opinion pollster soliciting federal voting preferences from a randomly selected Quebec voter. "A federal election has been called for October 19. Which party are you most likely to vote for in this election?" There are six possible options (Bloc Québécois, Conservative, Green, Liberal, New Democrat, or Other), and so we might consider a die where side i occurs with probability p_i satisfying $p_1 + p_2 + p_3 + p_4 + p_5 + p_6 = 1$.

Thus, if we ever want to model a chance experiment where there are finitely many possible outcomes, say N, then all we need to do is consider an N-sided die where side i appears with probability p_i satisfying $p_1 + \cdots + p_N = 1$. We can then take our sample space to be

$$S = \{1, 2, \ldots, N\},$$

and define the probability $\mathbf{P}$ by setting

$$\mathbf{P}(\{1\}) = p_1, \quad \mathbf{P}(\{2\}) = p_2, \quad \ldots, \quad \mathbf{P}(\{N\}) = p_N.$$

If we define the random variable X to equal the value of the outcome observed so that $X(i) = i$, then it follows that

$$\mathbf{P}(X = 1) = p_1, \quad \mathbf{P}(X = 2) = p_2, \quad \ldots, \quad \mathbf{P}(X = N) = p_N.$$