CONTENT OVERVIEW

Toss a coin or choose a simple random sample (SRS) from a population. The results can't be predicted in advance. When we flip a coin, we know we will get a head or a tail. We expect both outcomes to be equally likely, but we don't know for certain which outcome will occur the next time we flip the coin. An instructor chooses a random sample of four students each day to put homework problems on the board. Because the selection is random, each possible size-four sample is equally likely to be selected. One day Joe comes to class without having done his homework. If the class is small, his chances of getting chosen are pretty good. If his class is large, he is less likely to be in the selected sample. Joe won't know if he will be caught without his homework until the sample is actually drawn.

Flipping a coin and choosing a random sample are both examples of **random phenomena.** Other examples include: the outcome of rolling a die, the gender of the next person passing through a turnstile at a subway, the growth of a child in one month, the color of the next car that exits the parking lot, and whether guessing on a true-false question will result in a correct answer. In each of these cases, the next outcome is uncertain but, over the long run of many repetitions, a pattern emerges.

We use **probability** to assess the likelihood that a random phenomenon has a particular outcome. Probability is a number between 0 and 1. If A represents a particular outcome or set of outcomes of a random phenomenon, we write $P(A)$ to denote the probability that event A will occur. The closer $P(A)$ is to 0, the less likely it is for event A to occur. The closer $P(A)$ is to 1, the more likely it is for event A to occur. Probabilities are often expressed as percentages. For example, the probability of flipping two heads in a row is 0.25 or 25%.

Next, we look at three ways that probabilities can be assigned to events. First, suppose a sports reporter predicts that the Yankees have a 75% chance of beating the Red Sox in their next game. In this case, the reporter is most likely giving his or her professional assessment of the likelihood that the Yankees will win. That assessment is based on his or her knowledge of the players, whether the team is playing at their home field, past interactions between these two teams, and a whole host of other factors. So, we could classify this type of probability assignment as informed intuition.

Second, a large medical laboratory developed its own test for a person's vitamin D level. Too much vitamin D can be toxic, while insufficient vitamin D is linked to certain illnesses. After complaints that the test might be giving erroneous results, the company randomly selected a sample of patients and retested their vitamin D levels. Suppose that the sample consisted of 200 patients and that 18 of the initial tests were determined to be erroneous. The probability of an erroneous test result can be estimated from the proportion of erroneous tests found in the sample. In this case,

Probability of erroneous test = (frequency of erroneous test)/(number of tests)
= 18/200 = 0.09 or 9%.

Third, suppose a student did not study for a multiple choice exam. There were five choices for answers, (a) – (e), and only one correct answer. The student guessed the answer to each question without even reading the question. Here there is an underlying assumption that the student is equally likely to answer (a), (b), (c), (d), or (e) and that the same is true for the correct answer. To assess his probability of getting a correct answer, he used the probability formula for equally likely outcomes.

$$\text{Probability of an event} = \frac{\text{number of outcomes in the event}}{\text{total number of outcomes}}$$

$$\text{Probability of a correct answer} = \frac{\text{number of correct answers per question}}{\text{total number of ways to answer the question}} = \frac{1}{5} \text{ or } 20\%.$$

We have shown three ways to assign probabilities: informed intuition (educated guess), proportion (relative frequency), and a formula used when outcomes are equally likely. Remember that probabilities are always between 0 and 1. Anytime you calculate a probability and get something like 1.4 or -0.2, go back and check your calculations because you have made a mistake.