Why we need Probability?
Any time we want to answer a research question that involves using a sample to draw a conclusion about some larger population, we need to answer the question "how likely is it...?" or "what is the probability...?". To answer such a question, we need to understand probability, probability rules, and probability models.

Def. We call a phenomenon random if individual outcomes are uncertain but there is nonetheless a regular distribution of outcomes in a large number of repetitions.

Def. The probability of any outcome of a random phenomenon is the proportion of times the outcome would occur in a very long series of repetitions.

Example. Toss a coin [fair (balanced) coin]: 2 outcomes (H or T). Fig. 12.2 shows the results of tossing a coin 5000 times:
- as number of tosses increases, the proportion of Heads → 0.5;
- as number of tosses increases, the proportion of Tails also → 0.5.

+ Think about die example.

Figure 12.2. The proportion on tosses of a coin. This figure shows the results of 5000 tosses each.
Def. The **sample space** (S) of a random phenomenon is the set of all possible outcomes.

**Example.** For 1 coin toss \( S = \{H, T\} \).

**Example.** For 3 tosses of 1 coin / 3 coins:
\[
S = \{(H,H,H), (H,H,T), (H,T,H), (H,T,T),
(T,H,H), (T,H,T), (T,T,H), (T,T,T)\}
\]

**Example.** For 1 die toss (6-sided):
\[
S = \{1, 2, 3, 4, 5, 6\}.
\]

**Example.** For 2 dice: see Figure in Problem 3.

Def. **An event** is an outcome or set of outcomes of a random phenomenon.

What is a “set of outcomes”?

**Example.** For a die: toss 6 or 3, toss even number, …

**Example.** For coins: 2 Heads in a row, all three are the same, …

---

**Probability Rules**

1. The probability \( P(A) \) of any event \( A \) satisfies \( 0 \leq P(A) \leq 1 \).
2. If \( S \) is the sample space in a probability model, then \( P(S) = 1 \).
3. For any event \( A \), \( P(\text{A does not occur}) = 1 - P(A) \).
4. Two events \( A \) and \( B \) are **disjoint** if they have no outcomes in common and so can never occur simultaneously. If \( A \) and \( B \) are disjoint, \( P(\text{A or B}) = P(A) + P(B) \).

Def. A **finite probability model** – model with a finite sample space (= discrete probability models).

To assign probabilities in a finite model,

- list the probabilities of all the individual outcomes.
- These probabilities must be numbers between 0 and 1
- that add to exactly 1.
- The probability of any event is the sum of the probabilities of the outcomes making up the event.

**Example.**

<table>
<thead>
<tr>
<th>Days</th>
<th>0</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
</tr>
</thead>
<tbody>
<tr>
<td>Probability</td>
<td>0.73</td>
<td>0.06</td>
<td>0.06</td>
<td>0.06</td>
<td>0.04</td>
<td>0.02</td>
<td>0.01</td>
<td>0.02</td>
</tr>
</tbody>
</table>
Def. A **continuous probability model** assigns probabilities as areas under a density curve. The area under the curve and above any range of values is the probability of an outcome in that range.

Example.

Def. A **random variable** is a variable whose value is a numeric outcome.

---

**Problem 1.**
Probability is a measure of how likely an event is to occur. Match one of the probabilities that follow with each statement of likelihood given.

0 0.01 0.45 0.50 0.55 0.99 1

a) The event is impossible. It can never occur.

b) This event is certain. It will occur on every trial.

c) The event is very likely, but once in a while it will not occur in a long sequence of trials.

d) This event will occur slightly less often than not.

**Problem 2.**
Describe a sample space S for the random phenomenon:

a) A basketball player shoots 3 free throws. You record the sequence of hits and misses.

b) A basketball player shoots 3 free throws. You record the number of baskets he makes.

Problem 3.

Figure 12.2 from the book. The 36 possible outcomes in rolling two dice. Since the dice are fair, each outcome is equally likely. Each outcome has probability 1/36.

a) Find the probability of the event “roll 6”.

b) Find \( P(\text{outcome is even}) \).

Problem 4.

If you draw an M&M candy at random from a bag of the candies, the candy you draw will have one of the seven colors. The probability of drawing each color depends on the proportion of each color among all candies made. Here is the distribution for milk chocolate M&Ms:

<table>
<thead>
<tr>
<th>Color</th>
<th>Purple</th>
<th>Yellow</th>
<th>Red</th>
<th>Orange</th>
<th>Brown</th>
<th>Green</th>
<th>Blue</th>
</tr>
</thead>
<tbody>
<tr>
<td>Probability</td>
<td>0.2</td>
<td>0.2</td>
<td>0.2</td>
<td>0.1</td>
<td>0.1</td>
<td>0.1</td>
<td>?</td>
</tr>
</tbody>
</table>

a) What must be the probability of drawing a blue candy?

b) What is the probability that you do not draw a brown candy?

c) What is the probability that the candy you draw is either yellow, orange, or red?
Problem 5.
Michelle has a bag of colored candy-coated chocolates. The probabilities of each color are:

<table>
<thead>
<tr>
<th>Color</th>
<th>Probability</th>
</tr>
</thead>
<tbody>
<tr>
<td>Brown</td>
<td>0.15</td>
</tr>
<tr>
<td>Yellow</td>
<td>??</td>
</tr>
<tr>
<td>Red</td>
<td>0.10</td>
</tr>
<tr>
<td>Blue</td>
<td>??</td>
</tr>
<tr>
<td>Orange</td>
<td>0.20</td>
</tr>
<tr>
<td>Green</td>
<td>??</td>
</tr>
</tbody>
</table>

The probability of drawing a brown or a green candy is 0.30, and the probability of not drawing a yellow candy is 0.70. What is the probability of drawing a blue candy?

Problem 6.
Choose an adult in the United States at random and ask, "How many days per week do you lift weights?" Call the response X for short. Based on a large sample survey, here is a probability model for the answer you will get:

<table>
<thead>
<tr>
<th>Days</th>
<th>Probability</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>0.73</td>
</tr>
<tr>
<td>1</td>
<td>0.06</td>
</tr>
<tr>
<td>2</td>
<td>0.06</td>
</tr>
<tr>
<td>3</td>
<td>0.04</td>
</tr>
<tr>
<td>4</td>
<td>0.02</td>
</tr>
<tr>
<td>5</td>
<td>0.01</td>
</tr>
<tr>
<td>6</td>
<td>0.01</td>
</tr>
<tr>
<td>7</td>
<td>0.02</td>
</tr>
</tbody>
</table>

a) Verify that this is a legitimate finite probability model.

b) Describe the event $X < 4$ in words. What is $P(X < 4)$?

c) Express the event "lifted weights at least once" in terms of X. What is the probability of this event?

Problem 7.
Generate two random numbers between 0 and 1 and take X to be their sum. The sum X can take any value between 0 and 2. The density curve of X is the triangle shown below.

a) Verify by geometry that the area under the curve is 1.

b) What is the probability that X is less than 1?

c) What is the probability that X is less than 0.5?

Review how to deal with Uniform distributions!!!
**Problem 8.**
A couple plans to have three children. There are 8 possible arrangements of girls and boys. For example, GGB means the first two children are girls and the third child is a boy. All 8 arrangements are (approximately) equally likely.

a) Write down all 8 arrangements of the sexes of three children. What is the probability of any one of these arrangements?

b) Let X be the number of girls the couple has. What is the probability that X = 2?

c) Starting from your work in a), find the distribution of X. That is, what values can X take, and what are the probabilities for each value?

**Problem 9.**
Two special tetrahedral (four-sided) dice are rolled. On each die, each side is labeled with 1, 2, 3, or 4 dots. After each roll, the sum of the number of dots on the up-faced sides is recorded. Let E be the event that the sum is even. Let F be the event that the sum is 5 or more. Are E and F disjoint? Why or why not?

**Problem 10.**
Assume that event A occurs with probability 0.17 and event B occurs with probability 0.73. Assume that A and B are disjoint events. Which of the following must be true?

- a. It is possible that neither A nor B will occur.
- b. If A occurs, then B does not occur.
- c. The probability that A does not occur is 0.83.
- d. All of the above

**Problem 11.**
You randomly select 600 students and observe that 90 of them smoke. **Estimate** the probability that a randomly selected student smokes.

- a. 0.20
- b. 0.50, since there are two possible outcomes for every student surveyed (smoke, don’t smoke)
- c. 0.15
- d. 1.2
Problem 12.
We are interested in whether the drawn card (full deck of 52 cards) is red or black, also we have a fair 4-sided die.

a) If both the cards and die are tossed, using the letters B and R and the numbers 1 through 4, what is the sample space?

b) If the cards and die tossed, what is the probability the outcome will be Black card and even number?

c) Let A be the event the drawn card is black from the experiment above and B is the event that the die lands with an even number shown to us from the same experiment. Do events A and B are disjoint or not?