Our First Steps in Probability

Think about the following question:

Suppose we deal a 5-card hand from a regular 52-card deck. Which is larger, $P(1 \text{ king})$ or $P(2 \text{ hearts})$?

We will be able to solve this later. For now, we are building the foundation.
Example: Roll 2 dice, say a blue one and a yellow one.

Find \( P(\text{blue} + \text{yellow} = 6) \).

Possible outcomes (blue, then yellow):

\[
\begin{array}{ccccccc}
1,1 & 1,2 & 1,3 & 1,4 & 1,5 & 1,6 \\
2,1 & 2,2 & 2,3 & 2,4 & 2,5 & 2,6 \\
3,1 & 3,2 & 3,3 & 3,4 & 3,5 & 3,6 \\
4,1 & 4,2 & 4,3 & 4,4 & 4,5 & 4,6 \\
5,1 & 5,2 & 5,3 & 5,4 & 5,5 & 5,6 \\
6,1 & 6,2 & 6,3 & 6,4 & 6,5 & 6,6 \\
\end{array}
\]

All equally likely! (Very important.)

Now, which ones are we interested in?

\[
\begin{array}{ccccccc}
1,1 & 1,2 & 1,3 & 1,4 & 1,5 & 1,6 \\
2,1 & 2,2 & 2,3 & 2,4 & 2,5 & 2,6 \\
3,1 & 3,2 & 3,3 & 3,4 & 3,5 & 3,6 \\
4,1 & 4,2 & 4,3 & 4,4 & 4,5 & 4,6 \\
5,1 & 5,2 & 5,3 & 5,4 & 5,5 & 5,6 \\
6,1 & 6,2 & 6,3 & 6,4 & 6,5 & 6,6 \\
\end{array}
\]

5 out of 36 possible outcomes have blue+yellow = 6. So,

\[
P(\text{blue} + \text{yellow} = 6) = \frac{5}{36}
\]
But what does that really mean?

Here our “experiment” is to roll 2 dice.

Imagine doing the experiment many, many times, recording the results in a large notebook:

- Roll the dice the first time, and write the outcome on the first line of the notebook.
- Roll the dice the second time, and write the outcome on the second line of the notebook.
- Roll the dice the third time, and write the outcome on the third line of the notebook.
- Roll the dice the fourth time, and write the outcome on the fourth line of the notebook.
- Etc. Imagine you keep doing this, thousands of times, filling thousands of lines in the notebook.
The notebook might look like this:

<table>
<thead>
<tr>
<th>outcome</th>
<th>blue+yellow = 6?</th>
</tr>
</thead>
<tbody>
<tr>
<td>blue 2, yellow 6</td>
<td>no</td>
</tr>
<tr>
<td>blue 3, yellow 1</td>
<td>no</td>
</tr>
<tr>
<td>blue 1, yellow 1</td>
<td>no</td>
</tr>
<tr>
<td>blue 4, yellow 2</td>
<td>yes</td>
</tr>
<tr>
<td>blue 1, yellow 1</td>
<td>no</td>
</tr>
<tr>
<td>blue 3, yellow 4</td>
<td>no</td>
</tr>
<tr>
<td>blue 5, yellow 1</td>
<td>yes</td>
</tr>
<tr>
<td>blue 3, yellow 6</td>
<td>no</td>
</tr>
<tr>
<td>blue 2, yellow 5</td>
<td>no</td>
</tr>
<tr>
<td>...</td>
<td>...</td>
</tr>
<tr>
<td>...</td>
<td>...</td>
</tr>
<tr>
<td>...</td>
<td>...</td>
</tr>
</tbody>
</table>

Here 2/9 of the lines say “yes.”

But after many, many repetitions, approximately 5/36 of the lines will say “yes.”

For example, about how may lines will say “yes” if we do the experiment 720 times?
This is what probability really is: In what fraction of the lines does something of interest happen?

**It sounds simple, but if you always think about this “lines in the notebook” idea, probability problems are a lot easier to solve.**
So, \( P(\text{blue+yellow}=6) = 5/36. \)

What about \( P(\text{blue} = 2) \)?

You might answer, “It depends on what the yellow die is.” But that is wrong. You should reason as follows:

“If I do this experiment many, many times, approximately what fraction of the lines will have \( \text{blue} = 2 \)?”

<table>
<thead>
<tr>
<th>outcome</th>
<th>( \text{blue} = 2 )?</th>
</tr>
</thead>
<tbody>
<tr>
<td>blue 2, yellow 6</td>
<td>yes</td>
</tr>
<tr>
<td>blue 3, yellow 1</td>
<td>no</td>
</tr>
<tr>
<td>blue 1, yellow 1</td>
<td>no</td>
</tr>
<tr>
<td>blue 4, yellow 2</td>
<td>no</td>
</tr>
<tr>
<td>blue 1, yellow 1</td>
<td>no</td>
</tr>
<tr>
<td>blue 3, yellow 4</td>
<td>no</td>
</tr>
<tr>
<td>blue 5, yellow 1</td>
<td>no</td>
</tr>
<tr>
<td>blue 3, yellow 6</td>
<td>no</td>
</tr>
<tr>
<td>blue 2, yellow 5</td>
<td>yes</td>
</tr>
<tr>
<td>…</td>
<td>…</td>
</tr>
<tr>
<td>…</td>
<td>…</td>
</tr>
<tr>
<td>…</td>
<td>…</td>
</tr>
</tbody>
</table>
So, what fraction of lines will have say “yes”? 

<table>
<thead>
<tr>
<th>1,1</th>
<th>1,2</th>
<th>1,3</th>
<th>1,4</th>
<th>1,5</th>
<th>1,6</th>
</tr>
</thead>
<tbody>
<tr>
<td>2,1</td>
<td>2,2</td>
<td>2,3</td>
<td>2,4</td>
<td>2,5</td>
<td>2,6</td>
</tr>
<tr>
<td>3,1</td>
<td>3,2</td>
<td>3,3</td>
<td>3,4</td>
<td>3,5</td>
<td>3,6</td>
</tr>
<tr>
<td>4,1</td>
<td>4,2</td>
<td>4,3</td>
<td>4,4</td>
<td>4,5</td>
<td>4,6</td>
</tr>
<tr>
<td>5,1</td>
<td>5,2</td>
<td>5,3</td>
<td>5,4</td>
<td>5,5</td>
<td>5,6</td>
</tr>
<tr>
<td>6,1</td>
<td>6,2</td>
<td>6,3</td>
<td>6,4</td>
<td>6,5</td>
<td>6,6</td>
</tr>
</tbody>
</table>

6 of the 36 possibilities have blue = 2, and they are all equally likely, so about \( \frac{6}{36} = \frac{1}{6} \) of the lines will have blue = 2.

So, \( P(\text{blue} = 2) = \frac{1}{6} \).

Also:

\( P(\text{blue} = 3 \text{ and yellow} = 6) = \frac{1}{36} \).

\( P(\text{blue} = 3 \text{ or yellow} = 6) = \frac{11}{36} \).
Problems to be worked in class:

<table>
<thead>
<tr>
<th>1,1</th>
<th>1,2</th>
<th>1,3</th>
<th>1,4</th>
<th>1,5</th>
<th>1,6</th>
</tr>
</thead>
<tbody>
<tr>
<td>2,1</td>
<td>2,2</td>
<td>2,3</td>
<td>2,4</td>
<td>2,5</td>
<td>2,6</td>
</tr>
<tr>
<td>3,1</td>
<td>3,2</td>
<td>3,3</td>
<td>3,4</td>
<td>3,5</td>
<td>3,6</td>
</tr>
<tr>
<td>4,1</td>
<td>4,2</td>
<td>4,3</td>
<td>4,4</td>
<td>4,5</td>
<td>4,6</td>
</tr>
<tr>
<td>5,1</td>
<td>5,2</td>
<td>5,3</td>
<td>5,4</td>
<td>5,5</td>
<td>5,6</td>
</tr>
<tr>
<td>6,1</td>
<td>6,2</td>
<td>6,3</td>
<td>6,4</td>
<td>6,5</td>
<td>6,6</td>
</tr>
</tbody>
</table>

\[ P(\text{blue} + \text{yellow} = 7) = ? \]

\[ P(\text{blue} = 4 \text{ and yellow} = 1) = ? \]

\[ P(\text{blue} = 3 \text{ or yellow} = 3) = ? \]
Let’s move to something a bit more complicated:

The teacher asks Bill, Ann, Tom, Cathy and Linda to draw straws. The two kids who get the two shortest straws get a free ice cream cone.

P(Cathy get a cone) = ?

P(no boy gets a cone) = ?

P(one boy and one girl get a cone) = ?

Possible outcomes (be careful!):

B, A  B, T  B, C  B, L
A, T  A, C  A, L
T, C  T, L
C, L

(We don’t include, for example, A,B, since it’s the same to us as B,A. They both get cones.)

There are 10 possible outcomes, and all are equally likely. So:

P(Cathy get a cone) = 4/10 = 2/5

P(no boy gets a cone) = 3/10

P(one boy and one girl get a cone) = 6/10 = 3/5
### Outcome Cathy gets a cone?

<table>
<thead>
<tr>
<th>Outcome</th>
<th>Cathy gets a cone?</th>
</tr>
</thead>
<tbody>
<tr>
<td>B, T</td>
<td>no</td>
</tr>
<tr>
<td>T, L</td>
<td>no</td>
</tr>
<tr>
<td>C, L</td>
<td>yes</td>
</tr>
<tr>
<td>A, C</td>
<td>yes</td>
</tr>
<tr>
<td>A, T</td>
<td>no</td>
</tr>
<tr>
<td>...</td>
<td>...</td>
</tr>
<tr>
<td>...</td>
<td>...</td>
</tr>
</tbody>
</table>

After, say, 1,000 repetitions of this experiment, about how many lines will say “yes”?

\[
\frac{2}{5} \times 1,000 = 200 \text{ lines}
\]
Homework:

Main problem: Add David to the list of kids who will draw straws for the two ice cream cones. (So, there are now 6 kids in all.) Answer the same 3 questions as before: $P(\text{Cathy gets a cone}) = ?$, etc.

Challenge problem: Suppose our experiment is to roll 4 dice, not 2. Say we have one blue die, one yellow, one green, one red. Find $P(\text{blue+yellow+green+red} = 9)$. (There are over 1,000 possible outcomes of this experiment, compared to 36 for the 2-dice example above. Do NOT list all 1,000+ outcomes. Find a shortcut.)

Next week:

Conditional probability.