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 king) or P(2 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(blue + yellow = 6).

Possible outcomes (blue, then yellow):

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

All equally likely! (Very important.)

Now, which ones are we interested in?

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

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

$$P(blue + yellow = 6) = \frac{5}{36}$$

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:

```
    outcome                      blue+yellow = 6?

blue 2, yellow 6                       no
blue 3, yellow 1                       no
blue 1, yellow 1                       no
blue 4, yellow 2                       yes
blue 1, yellow 1                       no
blue 3, yellow 4                       no
blue 5, yellow 1                       yes
blue 3, yellow 6                       no
blue 2, yellow 5                       no
...                                    ...
...                                    ...
...                                    ...
```

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(blue+yellow=6) = 5/36.

What about P(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 blue = 2?"

```
    outcome                          blue = 2?

blue 2, yellow 6                       yes
blue 3, yellow 1                       no
blue 1, yellow 1                       no
blue 4, yellow 2                       no
blue 1, yellow 1                       no
blue 3, yellow 4                       no
blue 5, yellow 1                       no
blue 3, yellow 6                       no
blue 2, yellow 5                       yes
...                                    ...
...                                    ...
...                                    ...
```

So, what fraction of lines will have say “yes”?

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

6 of the 36 possibilities have blue = 2, and they are all equally likely, so about 6/36 = 1/6 of the lines will have blue = 2.

So, P(blue = 2) = 1/6.

Also:

P(blue = 3 and yellow = 6) = 1/36.

P(blue = 3 or yellow = 6) = 11/36.

Problems to be worked in class:

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

P(blue + yellow = 7) = ?

P(blue = 4 and yellow = 1) = ?

P(blue = 3 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?

B,T                             no
T,L                             no
C,L                             yes
A,C                             yes
A,T                             no
...                             ...
...                             ...
```

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

$\frac{2}{5} \times 1,000 = 200$ 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(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(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.