Activity 2: Introduction to Probability & Counting

1. In the last activity, we flipped coins under the assumption that we had a 50% chance of getting either heads or tails. If that assumption is true, why don't we expect to obtain exactly 50 heads if we toss a coin 100 times? How can we even be sure we have a 50% chance of flipping heads on any single coin toss?

> A phenomenon is **random** if individual outcomes are uncertain but there is nonetheless a regular distribution of outcomes in a large number of repetitions.
>
> The **probability** of any outcome refers to our prediction of the likelihood of that outcome occurring in the future.

2. What is the probability that I am reading this question on the second day of class? ____________________

   What is the probability that I will read this question on the second day of class next year? ____________________

   What is the probability that you showed up to class today? ____________________

> One of our primary goals for this first unit is to learn how to develop probability models.
>
> A **probability model** displays (graphically, via formulas, or through a list):
> - The **sample space** of an experiment (all possible outcomes of that experiment)
> - The **probability** of observing each of those outcomes

3. Write out the probability model for tossing a fair, six-sided die.

   Write out the probability model for your grade in this course. How did you estimate the probabilities for each model?

.

Probabilities are extremely useful, but they can be interpreted in many ways. Physical probabilities, such as those arising from coin tosses, card games, or radioactive atoms, can be interpreted and estimated a couple different ways:

Relative frequency approach: Probability is how frequently something happens over a large number of repetitions
Classical approach: Probability is related directly to the number of possible outcomes in an experiment.

Evidential probabilities, such as the our beliefs in the fairness of a coin or the probability that it will rain tomorrow, can be estimated through a subjective approach:

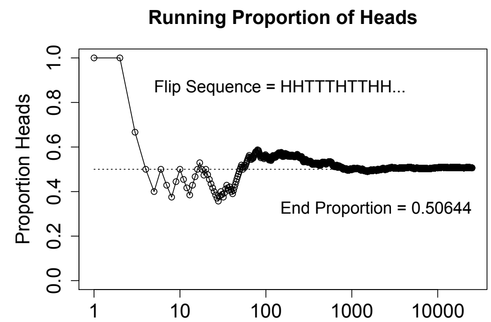
Subjective approach: Probability represents our degree of belief in an outcome based on previous evidence.

Relative Frequency Approach: $P(\text{win}) \approx \lim_{\text{trials} \to \infty} \frac{\text{wins}}{\text{trials}}$

We estimate probability as the proportion of times an outcome occurs in a long series of repetitions.

4. Suppose we toss a coin 10 times and observe 7 heads. If this was the only information we had about this coin, what would be our best estimate for the probability of tossing heads? Do we know this estimate is inaccurate? Explain.

   Best estimate for P(heads) = __________ Inaccurate?



5. Karl Pearson once tossed a coin 24,000 times and observed 12,012 heads. Using his results, we'd estimate the probability of tossing heads to be 0.5005.

   I had a computer simulate Pearson's experiment. The graph to the right shows the running proportion of heads when a simulated coin is flipped 25,000 times. While the outcome of each coin flip is random, the long-run relative frequency settled right near 0.50.

   Under the relative frequency approach, a greater number of repetitions yields a probability estimate that is more stable.

   Using the relative frequency approach, how would you estimate the following probabilities?

   a) What's the probability that the current vice president will become president in the future?

   c) If we roll two dice, what's the probability they sum to 7? To estimate this probability, I had a computer simulate 5,000 rolls of two dice. 897 of these 5000 rolls summed to 7. From this, what's our best estimate for the probability of rolling two dice and getting a sum of 7? Can you think of another way to estimate the probability of rolling a 7?

6. You are the world's laziest nurse. One night, four mothers with the last names Anderson, Boyd, Carter, and Dillon give birth to baby boys. Each mothers gives her child a first name alliterative to his last: Andy Anderson, Bobby Boyd, Chris Carter, and Dennis Dillon. You're too lazy to keep track of which baby belongs to which mother, so you decide to return the babies to their mothers completely at random. Use a simulation approach to estimate the probability you return 0, 1, 2, 3, or 4 babies to their correct mothers.

Trial #	**1**	**2**	**3**	**4**	**5**
# of correct matches:	_______	_______	_______	_______	_______

# of correct matches:	**0**	**1**	**2**	**3**	**4**	**Total**
proportion of trials:	_______	_______	_______	_______	_______	_______
actual probability (from ? #16):	_______	_______	_______	_______	_______	_______

Notice one potential problem with the relative frequency approach. Each time we collect (or simulate) more trials, our estimated probability will change. Perhaps we might want to use an approach that gives us a single "best" estimate for our probabilities of interest. This is the **Classical / Theoretical Approach**:

$$P\left(\text{event of interest}\right) = \frac{\text{\# of outcomes that result in the event of interest}}{\text{total number of possible outcomes}} = \frac{\text{\# of wins}}{\text{\# of possibilities}}$$

7. The following table shows the 36 possible outcomes from rolling 2 dice. Is each outcome equally likely? What's the probability of rolling a sum of 12? What's the probability of rolling a sum of 7? How does this probability compare to what we calculated under the relative frequency approach?

11	12	13	14	15	16
21	22	23	24	25	26
31	32	33	34	35	36
41	42	43	44	45	46
51	52	53	54	55	56
61	62	63	64	65	66

Each outcome equally-likely? **YES** **NO**

P(sum of 12) = ______________________________

Compared to relative frequency estimate of 0.1794

8. As you can see, in order to estimate probabilities using this classical approach, we must be able to **count** the total number of possible outcomes. Suppose we toss a coin 3 times. How many outcomes are possible? Is each outcome equally likely to occur? What's the probability that we observe at least 2 heads?

9. The classical approach is useful when estimating probabilities (**only if all outcomes are equally likely**), but counting outcomes can quickly become tedious. For example:

Experiment	Number of outcomes	
Rolling a 6-sided die 2 times	36	
Rolling a 6-sided die 10 times	60,466,176	
Tossing a coin 10 times	1,024	
Tossing a coin 20 times	1,048,576	
Randomly dividing 20 students into 4 groups of 5 students	488,864,376	(if group ID does not matter)
Randomly dividing 20 students into 4 groups of 5 students	11,732,745,024	(if group identification matters)
Choosing 6 lottery numbers between 1-36 with repeats	2,176,782,336	
Choosing 6 lottery numbers between 1-36 with no repeats	1,402,410,240	

Remember, to estimate probabilities using the classical approach, we need to be able to count all possible outcomes of an experiment (we do NOT, thankfully, need to list them all out).

To help count all possible outcomes in an experiment, we'll learn several counting rules. Because these rules are so useful, we will try to gain a deep understanding of them. When it comes to actually using the rules to calculate the number of outcomes, we'll use a calculator or computer.

By the end of this activity, I'm hoping you'll know how to calculate the number of outcomes in an experiment by using either combinations or something I call the "slot method." Let's see if we can derive these counting rules.

10. A local deli advertises a lunch special (sandwich, soup, dessert, and drink) for $5. They offer the following choices:
    - 5 sandwiches: chicken salad, ham, tuna, roast beef, and turkey
    - 4 soups: tomato, chicken noodle, vegetable, broccoli cheese
    - 2 desserts: cookie or cake
    - 5 drinks: water, tea, coffee, milk, juice

    How many different lunch specials are there? See if you can create a tree diagram to display this visually. Then, answer this question using the slot method.

    If you randomly choose a lunch special, what's the probability that it includes a ham sandwich?

11. Suppose you forget to study and you randomly guess on a 10-question true/false quiz. What's the probability that you get a perfect score? What's the probability you get at least one question correct?

> We just derived the **Fundamental Counting Principle (multiplication rule)**
>
> Suppose an experiment consists of a sequence of k choices. If there are n_1 choices in the first stage, n_2 choices in the second stage, and so on, the total number of outcomes in the experiment are:
>
> $$n_1 \times n_2 \times n_3 \times \ldots \times n_k$$

12. In order to create an online bank account, you must select a 4-character password. Each character in your password may be:
    - a lower-case letter
    - an upper-case letter
    - a number

    The characters may be repeated. If there are 300 million people in America, could they all have unique passwords?

13. How many passwords could we create if characters cannot be repeated (e.g., you cannot have A1Ag)?

14. How many passwords could we create if characters CAN be repeated but you cannot repeat the same character back-to-back (e.g., you can have A1Ag but not AA1g)?

15. My two older brothers and I all have the same initials, "BAT." According to babynamesworld.parentsconnect.com, there are:
    - 104 commonly used (English) boy names that begin with the letter "B."
    - 157 commonly used (English) boy names that begin with the letter "A."

    If my parents chose my first and middle names at random (ensuring I have the initials "BA"), what's the probability I would have gotten the name Bradley Adam? Can you think of a better way to estimate this probability?

16. You are once again the world's laziest nurse who has decided to randomly return 4 babies to 4 mothers. Calculate the number of possible outcomes in this experiment.

    Number of possible outcomes: ____

(4) ABCD	(2) BACD	(1) CABD	(0) DABC
(2) ABDC	(0) BADC	(0) CADB	(1) DACB
(2) ACBD	(1) BCAD	(2) CBAD	(1) DBAC
(1) ACDB	(0) BCDA	(1) CBDA	(2) DBCA
(1) ADBC	(0) BDAC	(0) CDAB	(0) DCAB
(2) ADCB	(1) BDCA	(0) CDBA	(0) DCBA

From this sample space, write out the probability model for this scenario (record these probabilities in question #6).

17. Suppose we randomly arrange 10 books on a shelf. What's the probability we (randomly) arrange them in alphabetical order?

> We've been using the **Factorial Rule**
>
> The number of ways to arrange n objects in a row, called a permutation of the n objects, is equal to
>
> $$n!$$
>
> We can only use this rule when we sample **without replacement**.

18. Suppose we decide to take a class photo. In how many ways could we line up for the photo?

19. Suppose we can only get 5 students in the class photo. How many different line-ups of 5 students could we select?

20. Using each digit only once, how many unique 4-digit numbers are there?

The last two questions are examples of **permutations**.

If we need to fill **r** positions by selecting from **n** different objects, we have ${}_nP_r = \dfrac{n!}{(n-r)!}$ arrangements

We can only use this rule when we sample **without replacement**.

We can only use this rule when **order does matter** (e.g., ABC is different from ACB)

21. How many ways are there to select a president, vice president, and secretary from 6 people? Is order important?

22. How many ways are there to select 3 committee members from 6 people? How does this differ from the previous question?

The last question was an example of a **combination**.

The number of ways to select r items from n distinct items (where order is not important) is

$${}_nC_r = \frac{n!}{(n-r)!r!}$$

We can only use this rule when we sample **without replacement**.

We can only use this rule when **order does not matter** (e.g., ABC is the same as ACB, BAC, BCA, CAB, CBA)

23. Will we always have fewer combinations than permutations? Why or why not?

24. How many different 5-card hands can we select from 52 cards? Use your calculator or a computer to get this answer.

25. Suppose we must assign 8 subjects to 2 groups (of 4 subjects each). In how many ways can we do this?

In how many ways can we assign 12 subjects into 3 groups (of 4 subjects each)?

Note: Your answers to the previous two questions are wrong. To see why, determine how many ways there are to assign 2 people to 2 groups.

26. In how many ways can we arrange the letters in the word MISSISSIPPI? Note that we're not distinguishing among the 4 S's, 4 I's, or 2 P's.

    Another way to ask this question would be: Out of 11 characters, in how many ways can we choose 1 M, 4 Is, 4 Ss, and 2 Ps?

27. Suppose we have 10 IE majors and 10 other majors in this class. I need to choose 4 students at random to fail this course (thus, satisfying my ego). In how many ways could I choose 4 students out of 20? In how many ways could I choose 4 students out of the 10 IE majors? Using the results from these two questions, what is the probability that I randomly select 4 IE majors to fail?