2 Working with Probabilities

- It is often helpful to think of probabilities as percentages or proportions.
- Furthermore, when working with multiple percentages, it is also helpful to construct hypothetical **two-way tables** (a.k.a., contingency tables) of counts.
- For the purposes of constructing the table and computing related probabilities, any value can be used for the hypothetical total count.
- When dealing with percentages (or proportions or probabilities) be sure to ask "percent *of what*?" Thinking in fraction terms, be careful to identify the correct reference group which corresponds to the denominator.

Example 2.1. Do American adults (18+) think it's acceptable to curse out loud in public? Assume[1] that

- 62% of American adults age 18-29 think it is acceptable
- 45% of American adults age 30-49 think it is acceptable
- 24% of American adults age 50-64 think it is acceptable
- 11% of American adults age 65+ think it is acceptable

Also assume that among American adults (18+)

- 20% of American adults are age 18-29
- 33% of American adults are age 30-49
- 25% of American adults are age 50-64
- 22% of American adults are age 65+

1. Consider a hypothetical group of 10000 American adults and assume the percentages provided apply to this group. Fill in the counts in each of the cells of the following table.

	18-29	30-49	50-64	65+	Total
Acceptable					
Not acceptable					
Total					10000

[1]The values in this problem are based on a March 12, 2025 article by the Pew Research Center.

2. Randomly select an American adult *from this group of 10000*. Compute the probability that they think cursing out loud in public is acceptable.

3. Randomly select an *American adult*. Compute the probability that they think cursing out loud in public is acceptable. (Hint: did the 10000 matter?)

4. Compute the probability that an American adult who thinks cursing out loud in public is acceptable is age 18-29. (Answer with both an unreduced fraction and a demical/percent.)

5. Compute the probability that an American adult who is age 18-29 thinks cursing out loud in public is acceptable. (Answer with both an unreduced fraction and a demical/percent.)

6. Compute the probability that an American adult is age 18-29 and thinks cursing out loud in public is acceptable. (Answer with both an unreduced fraction and a demical/percent.)

7. Compare the unreduced fractions for the previous three parts. What is the same? What is different?

8. Suppose that we were only told that 35.67% of American adults overall think cursing out loud in public is acceptable, and that we not given the values 62%, 45%, 24%, 11%. Would we be able to complete the two-way table?

- **Warning!** In general, knowing probabilities of individual events alone is not enough to determine probabilities of combinations of them.

Example 2.2. Suppose that 47% of American adults[2] have a pet dog and 25% have a pet cat.

1. Donny Don't says that 72% (which is 47% + 25%) of American adults have a pet dog or a pet cat. Is that necessarily true? If not, is it even possible (in principle anyway) for this to be true? Under what circumstance (however unrealistic) would this be true? Construct a corresponding two-way table.

2. Given only the information provided, what is the smallest possible percentage of American who adults have a pet dog or a pet cat? Under what circumstance (however unrealistic) would this be true? Construct a corresponding two-way table.

3. Donny Don't says that 11.75% (which is 47% $\times$ 25%) of Americans have both a pet dog *and* a pet cat. Explain to Donny why that's not necessarily true. Without further information, what can you say about the percent of American adults who have both a pet dog and a pet cat?

[2]These values are based on the 2018 General Social Survey.

4. Suppose that 14% of American adults have both a pet dog *and* a pet cat. What is the percentage of American adults who have a pet dog *or* a pet cat? Construct a corresponding two-way table. Use your table to show Donny how to correct his error from part 1.

5. What percentage of American adults who have a pet dog also have a pet cat? Is it 25%?

6. What percentage of American adults who do not have a pet dog have a pet cat? Is this the same value as in the previous part?

7. What percentage of American adults who have a pet cat also have a pet dog? Is it 47%?

8. Describe in words the percentage that results from subtracting the answer to the previous part from 100%.

Example 2.3. A woman's chances of giving birth to a child with Down syndrome increase with age. The CDC estimates[3] that a woman in her mid-to-late 30s has a risk of conceiving a child with Down syndrome of about 1 in 250. A nuchal translucency (NT) scan, which involves a blood draw from the mother and an ultrasound, is often performed around the 13th week of pregnancy to test for the presence of Down syndrome (among other things). If the baby has Down syndrome, the probability that the test is positive is about 0.9. However, when the baby does not have Down syndrome, there is still a probability that the test returns a (false) positive of about[4] 0.05. Suppose that the NT test for a pregnant woman in her mid-to-late 30s comes back positive for Down syndrome. What is the probability that the baby actually has Down syndrome?

1. Before proceeding, make a guess for the probability in question.

   0-20% 20-40% 40-60% 60-80% 80-100%

2. Donny Don't says: 0.90 and 0.05 should add up to 1, so there must be a typo in the problem. Do you agree?

3. Construct a hypothetical two-way table of counts to represent the given information.

4. Use the table to find the probability in question: If NT test for a pregnant woman in her mid-to-late 30s is positive, what is the probability that the baby actually has Down syndrome?

5. The probability in the previous part might seem very low to you. Explain why the probability is so low.

[3]Source: http://www.cdc.gov/ncbddd/birthdefects/downsyndrome/data.html

[4]Estimates of these probabilities vary between different sources. The values in the exercise were based on https://www.ncbi.nlm.nih.gov/pubmed/17350315

6. Compare the probability of having Down Syndrome before and after the positive test. How much more likely is a baby who tests positive to have Down Syndrome than a baby for whom no information about the test is available?

- Remember to ask "percentage *of what*"? For example, the percentage of *babies who have Down syndrome* that test positive is a very different quantity than the percentage of *babies who test positive* that have Down syndrome.
- Probabilities are often conditional on information.
- Conditional probabilities (e.g., probability of Down Syndrome *given a positive test*) can be highly influenced by the original unconditional probabilities (e.g. probability of Down Syndrome), sometimes called the **base rates**. Don't neglect the base rates when evaluating probabilities.
- The example illustrates that when the base rate for a condition is very low and the test for the condition is less than perfect there will be a relatively high probability that a positive test is a *false positive.*