2.1 Central Tendency

Dr. Lauren Perry

Goals

1. Calculate and interpret measures of center.
   - Learn the notation for a mean and for the sample size.
   - Calculate mean, median, and mode.
   - Understand what is meant by a "measure of center".
   - Determine which measure of center to use for a given dataset.

Measures of Central Tendency

Idea: What values are most common or most likely?

Three ways to measure:

- mode
- mean
- median

Mode

Mode: the most commonly occurring value.

- Used when working with categorical variables.
- We can get this easily from a frequency distribution.

Mean

Mean: this is what we usually think of as the "average".

- Denoted $\bar{x}$.
- Add up all of the values and divide by the number of observations (n):

$$\bar{x} = \frac{x_1 + x_2 + \cdots + x_n}{n} = \sum_{i=1}^{n} \frac{x_i}{n}$$

- x_i is the ith observation a
- $\sum_{i=1}^{n}$ is the sum of all observations from 1 through n.
  - This is called *summation notation*.

Median

Median: the middle number when the data are ordered from smallest to largest.

- If there are an odd number of observations, this will be the number in the middle:
  $\{1, 3, \mathbf{7}, 9, 9\}$ has median 7
- If there are an even number of observations, there will be two numbers in the middle. The median will be their average.
  $\{1, 2, \mathbf{4, 7}, 9, 9\}$ has median $\frac{4+7}{2} = 5.5$

The mean is sensitive to extreme values and skew. The median is not!

x: 1, 3, 7, 9, 9	y: 1, 3, 7, 9, 45
median $= 7$	median $= 7$
$\bar{x} = \frac{29}{5} = 5.8$	$\bar{y} = \frac{65}{5} = 13$

Changing that 9 out for a 45 changes the *mean* a lot! But the *median* is 7 for both x and y.

Because the median is not affected by extreme observations or skew, we say it is a **resistant measure** or that it is **robust**.

Which measure should we use?

- Mean: symmetric, numeric data
- Median: skewed, numeric data
- Mode: categorical data

Note: If the mean and median are roughly equal, it is reasonable to assume the distribution is roughly symmetric.

Weighted Means

Sometimes we have reason to calculate a **weighted mean**.

$$\bar{x}_w = w_1 x_1 + w_2 x_2 + \ldots w_n x_n$$

Each observation is multiplied by a corresponding *weight*, w.

- In general, each w represents the proportion attributed to that category.
- In general, we require that all of the w values sum to 1.

Weighted Means

If all the weights are equal, we get an (unweighted) mean:

- We would need n equal weights which sum to 1.
  - So each weight would be $1/n$

Here's what that looks like:

$$\frac{1}{n}x_1 + \frac{1}{n}x_2 + \cdots + \frac{1}{n}x_n = \frac{x_1 + x_2 + \cdots + x_n}{n} = \bar{x}$$

Example 1

Consider the following grade distribution:

- Assignments: 15%
- Quizzes: 20%
- Exam 1: 15%
- Exam 2: 15%
- Project: 15%
- Final Exam: 20%

We can use this to calculate an overall grade.

Example 1

Suppose some student has the following score in each category

- Assignments: 92%
- Quizzes: 76%
- Exam 1: 56%
- Exam 2: 69%
- Project: 89%
- Final Exam: 70%

Example 1

We can calculate their overall grade in the class using the weighted average formula.

- The proportion of the grade that comes from each category is the *weight*.
  - We will need to convert the percentages to proportions.
- The grades are our x values.

$$\begin{aligned}\text{grade} &= 92(0.15) + 76(0.20) + 56(0.15) + 69(0.15) + 89(0.15) + 70(0.20 \\ &= 13.8 + 15.2 + 8.4 + 10.35 + 13.35 + 14 \\ &= 75.1\end{aligned}$$

So this student would get a 75.1% in the class.

Example 2

Now suppose a student has the following scores

- Assignments: 83%
- Quizzes: 71%
- Exam 1: 61%
- Exam 2: 68%
- Project: 91%

and has not taken the final exam yet. He really wants to pass the class with at least a C-, but is not sure what kind of final exam grade would allow him to do that.

Example 2

- If he wants to pass, he needs a minimum overall grade of 70%.
  - (He needs his weighted average to be 70% or higher.)
- We know everything except his final exam score, so we'll make that F in our formula:

$$70 = 83(0.15) + 71(0.20) + 61(0.15) + 68(0.15) + 91(0.15) + F(0.20)$$

Example 2

To figure out what he needs to get on the final, we need to solve for F.

$$
\begin{aligned}
70 &= 83(0.15) + 71(0.20) + 61(0.15) + 68(0.15) + 91(0.15) + F(0.20) \\
70 &= 12.45 + 14.2 + 9.15 + 10.2 + 13.65 + 0.2F \\
70 &= 59.65 + 0.2F \\
10.35 &= 0.2F \\
F &= 51.75
\end{aligned}
$$

So he needs to get at least a 51.75% on the final exam in order to pass the class.