Measures of Center

Summary

1. Measures of center give us a "starting point" for our data set.

The Mean

Mean

The **mean** of a data set is found by adding all of the data values together and then dividing by the total number of values.

When most people use the term *average*, they are referring to the mean.

Properties of the Mean

- Sample means drawn from the same population tend to vary less than other measures of center.
- The mean of a data set uses every value, unless the mean is a *trimmed mean*.
- One extreme value (called an **outlier**) can change the value of the mean drastically.

Mean Formula

Sample mean: $\overline{x}$ Population mean: μ

$$\overline{x}, \text{ or } \mu, = \frac{\sum x_i}{n} = \frac{1}{n}\sum x_i$$

The Median

Median

The **median** of a data set is found by first arranging the data values from least to greatest, then selecting the data value in the middle.

Properties of the Median

- Denoted by $\tilde{x}$ or `Med`.
- If there are 2 data values in the middle, the median will be the mean of these two values.
- Separates the top 50% of the data from the bottom 50%.
- Typically does not change by large amounts when including extreme values (median is **resistant**)

The Mode

Mode

The **mode** of a data set is the value(s) that occur the most.

- May be one mode, no mode, or many modes (2 modes is called **bimodal**).
- Only measure of center to use on qualitative data.

Example 1. The data set below represents the number of complaints I receive each week about my teaching.

8, 2, 3, 7, 4, 4, 1, 9, 7, 5

Calculate the mean, median, and mode of the number of complaints.

Example 2. The next week, I received 400 complaints. Re-calculate the mean, median, and mode.

Example 3. California has a mean class size of 20.9 students per teacher and Alaska has a mean of 16.8 students per teacher.

Combining the two states, the mean number of students per teacher to be 18.85, $\left(\frac{20.9+16.8}{2}\right)$, but is this result correct? Why or why not?

Weighted Mean

Sometimes it is necessary to take into account how large each class of a data set is.

If w_i represents the **weight** of each class, then the **weighted mean** can be found via

$$\frac{\sum (x_i \cdot w_i)}{\sum w_i}$$

In other words:

1. Multiply each data value by its corresponding weight.
2. Add those results.
3. Then divide that by the total of the weights.

Example 4. You've recently completed a semester in college. Determine the semester's GPA (A = 4pts, B = 3pts, etc).

Course	Grade	Credit Hours
Statistics	A	4
Advanced Chris Farley	A	3
Airplane! Quotes	B	5
Obnoxious Examples	C	3

Example 5. In a statistics course, tests count for 60% of the final grade, homework for 20% and midterm and final exams are 10% each. Suppose you've earned an 87% average on tests, 94% average on homework and a 77% average on the exams.

What is your overall percentage for the course?

Grouped Data

- In grouped form, we don't know individual values in our data set.
- Mean can just be an educated guess.
- Use class midpoints and weighted mean techniques.

Example 6. The table below gives the number of sushi rolls various tables ordered at a Japanese restaurant one day.

Number of Rolls	Frequency
1 – 5	4
6 – 10	17
11 – 15	12

Estimate the mean number of sushi rolls consumed per table.