PSC 348: Political Analysis

November 2, 2004
Chris Cooper
Assistant Professor
Western Carolina University

Reviewing Last Class
- Order of operations
  - Parentheses
  - Exponents
  - Multiplication and division
  - Addition and subtraction

Data Presentation
- Measures of Central Tendency
  - Mean (Sum of all scores/Total # of observations)
    - Sensitive to outliers
  - Median (value dividing a distribution into two groups of equal size)
    - Less sensitive to outliers
  - Mode (Category of a variable having the largest number of instances)
    - Used mostly with nominal level data (PID)

Measures of Central Tendency
- Most common value = mode
  - Can be used on any level of measurement
  - To compute the mode:
    - Sort cases into groups
    - Count cases in each group
    - Mode = group with most cases
  - Limitations
    - may be more than one mode
    - Mode may not be representative

Measures of Central Tendency
- Mean is the arithmetic average (x bar)
  - Cannot be calculated for nominal variables
  - To calculate:
    - Mean = \( \frac{\sum x}{N} \)
  - Limitations: may give you misleading answers for skewed distributions

Measures of Central Tendency
- Median
  - Value which divides a set of observations in half
  - Cannot be calculated for nominal variables
  - How to calculate:
    - Arrange from lowest to highest
    - Take middle value if N is odd
    - If N is even, take the mean of the two middle values
Measures of Dispersion

- Quantifies the amount of variability in a given variable
  - Range
  - Interquartile Range
  - Standard Deviation
  - Variance

What are measures of dispersion?

- A MEASURE OF DISPERSION quantifies the amount of variability in a given variable

Range

- The RANGE is the difference between the highest value and the lowest value in a distribution.

Weaknesses of range

- The range does not take into account the values of all of the cases in a distribution
  - Example:
    - 0, 0, 0, 50, 100, 100, 100
    - 0, 0, 25, 50, 75, 100, 100

Interquartile Range

- Specifies the range of values within the middle 50 percent of the observations found

Example of Interquartile Range

- Eight salaries reported
- Omit the lowest 25% and the highest 25%
- Middle 50% range from 38,000 to 45,000
- IQR is 45,000 - 38,000 = 7,000
### Standard Deviation

- A measure of the average distance of values in a distribution from the arithmetic mean of the distribution
  - If all values are the same, then the standard deviation is 0
  - As the observations become more different from one another, the standard deviation becomes larger

### To calculate the standard deviation

- List each case
- Find the mean of the distribution
- Subtract the mean from each case
- Square it (for each case)
- Add them all together
- Divide by N (unless you have a sample, then divide by N-1)
- Take the square root of it all

### Variance

- The square of the standard deviation
- In other words, go through the same computation, just don’t take the square root.

\[
\text{variance} = \frac{\sum (X_i - \bar{X})^2}{N}
\]

### How to do it in .spss

- Click on:
  - Analyze
  - Descriptive Statistics
  - Frequencies
  - Click on the variable you want, then the right arrow
  - Statistics
  - Click on what you want
  - Continue
  - OK

### Wrapping up

- Read Chapter 13
- What was the most important point?
- What would you like to know more about?
- What was the muddiest point?