Health Science 503
Advanced Community Health Statistics
Lecture 2
Displays and Summaries of Data

Topics for Tonight
- Chapter 3: Organization & Presentation of Data
  - Frequency distributions
  - Graphic Displays
- Chapter 4: Summary Statistics
  - Measures of central tendency
  - Measures of dispersion

Frequency Distribution

<table>
<thead>
<tr>
<th>Score</th>
<th>Frequency</th>
<th>Percent</th>
<th>Cumulative Percent</th>
</tr>
</thead>
<tbody>
<tr>
<td>90-100</td>
<td>10</td>
<td>50.0</td>
<td>50.0</td>
</tr>
<tr>
<td>80-89</td>
<td>5</td>
<td>25.0</td>
<td>75.0</td>
</tr>
<tr>
<td>70-79</td>
<td>2</td>
<td>10.0</td>
<td>85.0</td>
</tr>
<tr>
<td>60-69</td>
<td>2</td>
<td>10.0</td>
<td>95.0</td>
</tr>
<tr>
<td>Below 60</td>
<td>1</td>
<td>5.0</td>
<td>100.0</td>
</tr>
<tr>
<td>Total</td>
<td>20</td>
<td>100.0</td>
<td></td>
</tr>
</tbody>
</table>
Chapter 4
Summary Statistics

- Measures of Central Tendency
  - Best single value to represent distribution of scores
  - Major and minor measures of central tendency
- Measures of Dispersion
  - How spread out are the scores
  - Major and minor measures of dispersion
Major Measures of Central Tendency
- Mean
- Median
- Mode

Arithmetic Mean or Average
- Sum of scores/ # of scores
- Two mean are of interest in this class
  - Population: \( \mu = \frac{\sum x_i}{N} \)
  - Sample: \( \bar{x} = \frac{\sum x_i}{n} \)

Median
- Median: Half of the scores less than this value
- Responsible for method when no ties
- Two cases:
  - Odd # of scores (middle score)
    - Scores sorted in ascending order
    - E.g., 1, 2, 5, 20, 500  Median = 5
  - Even # of scores (mean of two middle scores)
    - Scores sorted in ascending order
    - E.g., 1, 2, 4, 5, 7, 21  Median = (4+5)/2 = 4.5
Mode

- Mode is the most frequently occurring value
- E.g., 0, 0, 1, 2, 2, 2, 2, 2, 50, 91
- Mode = 2

Minor Measures of Central Tendency

- You are not responsible for the minor measures.
- What are the minor measures of central tendency?
  - Geometric mean
  - Harmonic mean

Geometric Mean

\[ GM = \sqrt[3]{x_1 \times x_2 \times \cdots \times x_n} \]

For example:

\[ GM = \sqrt[3]{2 \times 4 \times 1} = \sqrt[3]{8} = 2 \]
Harmonic Means

\[ HM = \frac{1}{\frac{1}{x_1} + \frac{1}{x_2} + \ldots + \frac{1}{x_n}} \]

For example: \( X = \{1, 2, 4, 8\} \)

\[ HM = \frac{1}{\frac{1}{1} + \frac{1}{2} + \frac{1}{4} + \frac{1}{8}} = \frac{1}{1.875} = 2.1333 \]

Relations Between Measures of Central Tendency

Mode, Median, & Mean

Relationship When Positive or Negative Skewness
Selecting a Measure of Central Tendency

- Data represent categories
  - Number of students in colleges on campus
  - Mode
- Highly skewed numeric data
  - Income
  - Median
- Fairly symmetric data
  - Scores on standardized tests
  - Mean

What Is Dispersion?

- Dispersion is how spread out the scores are.
- Dispersion can be important
  - Baker and Long Beach both have a mean temperature of 70 °F
  - Long Beach 40 °F To 100 °F
  - Baker 20 °F To 120 °F
- Minor and major measures of dispersion

Minor Measures of Dispersion

- Range = Maximum – Minimum
- Mean Absolute Deviation (MAD)

\[
MAD = \frac{\sum |X_i - \mu|}{N}
\]

\[
MAD = \frac{\sum |X_i - \mu|}{n}
\]
What Is Dispersion?

- Dispersion is how spread out the scores are.
- Dispersion can be important
  - Baker and Long Beach both have a mean temperature of 70°F
  - Long Beach 40°F To 100°F
  - Baker 20°F To 120°F
- Minor and major measures of dispersion
- Major measures of dispersion can be calculated for
  - Population
  - Sample

Major Measure of Dispersion for Population

Population Variance

\[ \sigma^2 = \frac{\sum (X - \mu)^2}{N} \]

Standard Deviation

\[ \sigma = \sqrt{\sigma^2} \]

Major Measure of Dispersion for Sample

Sample Variance

\[ S^2 = \frac{\sum (X - \overline{X})^2}{n-1} \]

Standard Deviation

\[ S = \sqrt{S^2} \]

Note: We are skipping grouped data calculations.
Introduction to SPSS

- SPSS is easy to use and very powerful
  - Most procedures are menu driven
  - User interface has number of windows
    - Data Window (In data mode similar to Excel)
    - Output Window
    - Syntax Window can be used to store commands

- This session
  - Enter data
  - Generate frequency distributions
  - Calculate mean and standard deviation