**MEASURES OF CENTRAL TENDENCY & VARIABILITY**

Prepared by: Jess Roel Q. Pesole

### COMPARISON

**CENTRAL TENDENCY**
- what is average or typical in a distribution
  - Commonly quantified characteristics of a distribution
  - Measures:
    1. Mode
    2. Median
    3. Mean

**VARIABILITY**
- extent to which scores are spread out in a distribution
  - Measures:
    1. Range
    2. Standard deviation
    3. Variance

### MEASURE OF CENTRAL TENDENCY: MODE (Mo)
- The most frequent score in the distribution
- For *ungrouped* data, it is the score or category that occurs most often in a distribution

Example (for ungrouped data):
1, 2, 3, 4, 4, 5, 6, 6, 7, 8, 9, 10

Mo = 6

### MEASURE OF CENTRAL TENDENCY: MODE (Mo)
- For grouped data, it is the midpoint of the interval containing the largest number of cases

Example:

<table>
<thead>
<tr>
<th>Class Interval</th>
<th>f</th>
</tr>
</thead>
<tbody>
<tr>
<td>60 – 64</td>
<td>12</td>
</tr>
<tr>
<td>55 – 59</td>
<td>17</td>
</tr>
<tr>
<td>50 – 54</td>
<td>28</td>
</tr>
<tr>
<td>45 – 49</td>
<td>30</td>
</tr>
<tr>
<td>40 – 44</td>
<td><strong>35</strong></td>
</tr>
<tr>
<td>35 – 39</td>
<td>18</td>
</tr>
<tr>
<td>30 – 34</td>
<td>10</td>
</tr>
</tbody>
</table>

Mo = (44 + 40) / 2 = 42
Therefore, Mo = 42.

### EXERCISE

- Find the mode of the following series of numbers.
  1. 24, 21, 20, 19, 18, 17, 16, 15, 14, 12
  2. 22, 20, 19, 18, 18, 18, 18, 16, 16, 16
  3. 28, 27, 27, 26, 26, 26, 26, 24, 24, 22

![Unimodal and bimodal histograms.](image)
**EXERCISE**

- Determine the mode for the following frequency distribution.

<table>
<thead>
<tr>
<th>Class Interval</th>
<th>f</th>
<th>cf</th>
<th>c%</th>
</tr>
</thead>
<tbody>
<tr>
<td>30 – 35</td>
<td>6</td>
<td>50</td>
<td>100.00</td>
</tr>
<tr>
<td>24 – 29</td>
<td>15</td>
<td>44</td>
<td>88.00</td>
</tr>
<tr>
<td>18 – 23</td>
<td>12</td>
<td>29</td>
<td>58.00</td>
</tr>
<tr>
<td>12 – 17</td>
<td>10</td>
<td>17</td>
<td>34.00</td>
</tr>
<tr>
<td>6 – 11</td>
<td>7</td>
<td>7</td>
<td>14.00</td>
</tr>
<tr>
<td><strong>N = 50</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**MEASURE OF CENTRAL TENDENCY:**  
**MEDIAN (Mdn)**

- Scale value below which 50 percent of scores fall
- Similar to $P_{50}$
- For **ungrouped** data, it is the centermost score if the number of scores are **odd**
  - If it is **even**, it is the average of the two centermost scores
- Position of the median = $(N + 1) / 2$

**Example (for ungrouped scores):**

11, 12, 13, 16, 17, 20, 25

- **N = 7** (odd)
- Therefore, the median is 16

11, 12, 13, 16, 17, 20, 25, 26

- **N = 8**
- $Mdn = \frac{(16 + 17)}{2} = 16.5$
- Therefore, the median is 16.5

**EXERCISE**

- Find the median of the following series of numbers.
  1. 24, 21, 20, 19, 18, 17, 16, 15, 14, 12
  2. 22, 20, 19, 18, 18, 18, 18, 16, 16, 10
  3. 28, 27, 26, 26, 26, 26, 24, 24, 22

**EXERCISE**

- Determine the median for the following frequency distribution.

<table>
<thead>
<tr>
<th>Class Interval</th>
<th>f</th>
<th>cf</th>
<th>c%</th>
</tr>
</thead>
<tbody>
<tr>
<td>30 – 35</td>
<td>6</td>
<td>50</td>
<td>100.00</td>
</tr>
<tr>
<td>24 – 29</td>
<td>15</td>
<td>44</td>
<td>88.00</td>
</tr>
<tr>
<td>18 – 23</td>
<td>12</td>
<td>29</td>
<td>58.00</td>
</tr>
<tr>
<td>12 – 17</td>
<td>10</td>
<td>17</td>
<td>34.00</td>
</tr>
<tr>
<td>6 – 11</td>
<td>7</td>
<td>7</td>
<td>14.00</td>
</tr>
<tr>
<td><strong>N = 50</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
MEASURE OF CENTRAL TENDENCY: MEAN

• Sum of the scores divided by the number of scores

\[ \bar{X} = \frac{\sum X_i}{N} \]

Where:
- \( \bar{X} \) = mean of a sample
- \( \mu \) = mean of a population
- \( \sum \) = summation
- \( X_1 \ldots X_N \) = raw scores
- \( N \) = number of scores

\[ \mu = \frac{\sum X_i}{N} \]

Example:
The following are scores from a sample of 10 math scores. Solve for the mean of the scores.
11, 12, 13, 14, 16, 17, 17, 20, 25, 26

\[ \bar{X} = \frac{11 + 12 + 13 + 14 + 16 + 17 + 17 + 20 + 25 + 26}{10} = 17.1 \]

Conclusion: The mean from the sample of math scores is 17.1.

MEASURE OF CENTRAL TENDENCY: MEAN

• If you are given grouped data:

\[ \bar{X} = \frac{\sum f m}{N} \]

Where:
- \( \bar{X} \) = mean of a sample
- \( \mu \) = mean of a population
- \( \sum \) = summation
- \( f \) = frequency of the interval
- \( m \) = midpoint of the interval
- \( N \) = number of scores

\[ \mu = \frac{\sum f m}{N} \]

Example:
The following are productivity scores for four departments in a company, as well as its corresponding number of workers. Compute for the overall mean productivity for all four departments:

- Section 1: \( N_1 = 20; \bar{X}_1 = 10 \)
- Section 2: \( N_2 = 15; \bar{X}_2 = 14 \)
- Section 3: \( N_3 = 18; \bar{X}_3 = 15 \)
- Section 4: \( N_4 = 22; \bar{X}_4 = 8 \)

EXERCISE
MEASURE OF CENTRAL TENDENCY: MEAN

• Deviation – distance and direction of any raw score from the mean

\[ \text{Deviation} = X - \bar{X} \]

Example: 9, 8, 6, 5, 2

\[
\begin{array}{c|c}
X & \text{DEVIATION} \\
9 & 3 \\
8 & 2 \\
6 & 0 \\
5 & 1 \\
2 & -4 \\
\end{array}
\]

MEASURES OF CENTRAL TENDENCY AND SYMMETRY

WHICH TO USE?

<table>
<thead>
<tr>
<th>MODE</th>
<th>Used for any measurement scale, especially the nominal scale</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Used when haste is necessary</td>
</tr>
<tr>
<td>MODE</td>
<td>Used for ordinal or interval data</td>
</tr>
<tr>
<td></td>
<td>Used when data is skewed, since it is not as sensitive to extreme scores compared to the mean</td>
</tr>
<tr>
<td>MEDIAN</td>
<td>- Used for interval or ratio data</td>
</tr>
<tr>
<td></td>
<td>- Used when one has a symmetrical or normal distribution</td>
</tr>
<tr>
<td></td>
<td>- Can also be useful in skewed distributions, since it is more flexible to advanced statistical analysis</td>
</tr>
</tbody>
</table>

PROPERTIES OF THE MEAN

1. The mean is sensitive to the exact value of all the scores in the distribution.
2. The sum of the deviations about the mean equals zero. \( \sum (X_i - \bar{X}) = 0 \)
3. The mean is very sensitive to extreme scores.
4. The sum of the squared deviations of all the scores about their mean is a minimum.
5. Under most circumstances, of the measures used for central tendency, the mean is least subject to sampling variation.

PROPERTIES OF THE MEDIAN

1. The median is less sensitive than the mean to extreme scores.
2. Under usual circumstances, the median is more subject to sampling variability than the mean but less subject to sampling variability than the mode.
MEASURES OF VARIABILITY: RANGE

- **Range** - Difference between the highest and lowest scores in a distribution

  \[ \text{Range} = \text{Highest score} - \text{lowest score} \]

  **Example:**
  
  4 17 12 9 6 10 1 5 9 3
  
  Range = 17 – 1
  
  = 16

MEASURES OF VARIABILITY: STANDARD DEVIATION

- **Measure of variability that reflects the typical deviation from the mean**

  - **Population Standard Deviation**
    \[ \sigma = \sqrt{\frac{\sum (X - \mu)^2}{N}} \]
  
  - **Sample Standard Deviation**
    \[ s = \sqrt{\frac{\sum (X - \bar{X})^2}{N - 1}} \]

  \( SS = \text{sum of squared deviations} \)
  
  \( \sigma = \text{population standard deviation} \)
  
  \( s = \text{sample standard deviation} \)

  \( \sum (X - \bar{X})^2 \)

  (estimated population standard deviation)

MEASURES OF VARIABILITY: STANDARD DEVIATION

- **The deviation equation (Deviation = X - \bar{X}) was only limited to only two values**

- **Adding all deviations in a distribution would only lead to ZERO.**

- **In the past, the mean deviation was used to determine the variability of a distribution. In this process, absolute deviations were used:**

  \[ MD = \frac{\sum |X - \bar{X}|}{N} \]

  These days, the mean deviation is no longer used since it is hard to use in advanced statistical analysis.

MEASURES OF VARIABILITY: STANDARD DEVIATION

- **However, the variance has the weakness of expressing variability in squared units**

  (EXAMPLE: If you were asked to determine the variability of exam scores in a class, the variance will express it in terms of squared exam score.)

- **As such, we get the square root of the deviance in order to reflect variability in the appropriate units; this becomes our sum of squares (SS)**

MEASURES OF VARIABILITY: STANDARD DEVIATION

- **Raw Score Formula for Standard Deviation:**

  \[ \sigma = \sqrt{\frac{\sum X^2}{N} - \bar{X}^2} \]

  \( s = \sqrt{\frac{\sum X^2}{N - 1} - \bar{X}^2} \)

  \( \sum X^2 = \text{sum of the squared raw scores} \)

  \( N = \text{total number of scores} \)

  \( \bar{X}^2 = \text{mean square} \)
MEASURES OF VARIABILITY:

STANDARD DEVIATION

• Example:
On a measure of authoritarianism (higher scores reflect greater tendency toward prejudice, ethnocentrism and submission to authority), seven students scored as follows:

<table>
<thead>
<tr>
<th>X</th>
<th>X²</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>6</td>
<td>36</td>
</tr>
<tr>
<td>6</td>
<td>36</td>
</tr>
<tr>
<td>7</td>
<td>49</td>
</tr>
<tr>
<td>4</td>
<td>16</td>
</tr>
<tr>
<td>10</td>
<td>100</td>
</tr>
</tbody>
</table>

\[ \Sigma X = 37 \quad \Sigma X^2 = 247 \]
\[ N = 7 \quad \text{Mean} = 5.2857 \]
\[ \text{Mean Square} = 27.9386 \]

PROPERTIES OF THE STANDARD DEVIATION

1. It gives a measure of dispersion relative to the mean.
2. It is sensitive to each score in the distribution.
3. It is stable with regards to sampling fluctuations.

MEASURES OF VARIABILITY:

VARIANCE

• Square of the standard deviation
\[ \sigma^2 = \frac{SS_{\text{pop}}}{N} \]
\[ S^2 = \frac{SS_{\text{pop}}}{N - 1} \]

\[ SS = \text{sum of squared deviations} \]
\[ \sigma = \text{population standard deviation} \]
\[ s = \text{sample standard deviation} \]
\[ s^2 = \text{sample standard deviation} \]

SUMMARY

<table>
<thead>
<tr>
<th>RAW SCORES</th>
<th>UNGROUPED FREQUENCY DISTRIBUTION</th>
<th>GROUPED FREQUENCY DISTRIBUTION</th>
</tr>
</thead>
<tbody>
<tr>
<td>MODE</td>
<td>MEDIAN</td>
<td>MEDIAN</td>
</tr>
</tbody>
</table>
| Score with the greatest frequency | Middlemost score, as determined by the position of the median | Mdn (50th percentile)= \[ X_L + \left( \frac{i}{f_L} \right) (\text{cum } f_r - \text{cum } f_L) \]

MEAN

<table>
<thead>
<tr>
<th>RAW SCORES</th>
<th>UNGROUPED FREQUENCY DISTRIBUTION</th>
<th>GROUPED FREQUENCY DISTRIBUTION</th>
</tr>
</thead>
<tbody>
<tr>
<td>[ \bar{X} \quad \mu = \frac{\sum X_i}{N} ]</td>
<td>[ \bar{X} \quad \mu = \frac{\sum fX}{N} ]</td>
<td>[ \bar{X} \quad \mu = \frac{\sum fm}{N} ]</td>
</tr>
</tbody>
</table>
### PRACTICE

The scores of attitudes toward older people for 30 students were arranged in the following simple frequency distribution (higher scores indicate more favorable attitudes towards older people):

<table>
<thead>
<tr>
<th>Attitude Score Value</th>
<th>( f )</th>
</tr>
</thead>
<tbody>
<tr>
<td>7</td>
<td>3</td>
</tr>
<tr>
<td>6</td>
<td>4</td>
</tr>
<tr>
<td>5</td>
<td>6</td>
</tr>
<tr>
<td>4</td>
<td>7</td>
</tr>
<tr>
<td>3</td>
<td>5</td>
</tr>
<tr>
<td>2</td>
<td>4</td>
</tr>
<tr>
<td>1</td>
<td>1</td>
</tr>
</tbody>
</table>

\( N = 30 \)

Find the (a) mode, (b) median, and (c) mean.

### PRACTICE

The scores on a religiosity scale (higher scores indicate greater commitment to religious expression) were obtained for 46 adults. For the following simple frequency distribution, calculate the three measures of central tendency.

<table>
<thead>
<tr>
<th>Score Value</th>
<th>( f )</th>
<th>( cf )</th>
</tr>
</thead>
<tbody>
<tr>
<td>10</td>
<td>3</td>
<td>46</td>
</tr>
<tr>
<td>9</td>
<td>4</td>
<td>43</td>
</tr>
<tr>
<td>8</td>
<td>6</td>
<td>39</td>
</tr>
<tr>
<td>7</td>
<td>8</td>
<td>33</td>
</tr>
<tr>
<td>6</td>
<td>9</td>
<td>25</td>
</tr>
<tr>
<td>5</td>
<td>7</td>
<td>16</td>
</tr>
<tr>
<td>4</td>
<td>5</td>
<td>9</td>
</tr>
<tr>
<td>3</td>
<td>2</td>
<td>4</td>
</tr>
<tr>
<td>2</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
</tbody>
</table>

\( N = 46 \)

35. A social psychologist interested in the dating habits of college undergraduates samples 10 students and determines the number of dates they have had in the last month. Given the scores 1, 8, 12, 3, 8, 14, 4, 5, 8, 16, compute the following:

a. Mean  
b. Median  
c. Mode  
d. Range  
e. Standard deviation  

### Summary

<table>
<thead>
<tr>
<th>STANDARD DEVIATION ( (\sigma \text{ or } \sigma') )</th>
<th>VARIANCE ( (\sigma^2 \text{ or } \sigma'^2) )</th>
</tr>
</thead>
</table>
| RAW SCORES                                       | \[
\sqrt{\frac{\sum X^2}{N - 1} - \bar{X}^2} \]
| UNGROUPED FREQUENCY DISTRIBUTION                  | \[
\frac{\sum fX^2}{N - 1} - \bar{X}^2 \]
| GROUPED FREQUENCY DISTRIBUTION                    | \[
\frac{\sum fm^2}{N - 1} - \bar{X}^2 \]

33. Without actually calculating the variability, study the following sample distributions:

Distribution a: 21, 24, 28, 22, 20  
Distribution b: 21, 32, 38, 15, 11  
Distribution c: 22, 22, 22, 22, 22  

a. Rank-order them according to your best guess of their relative variability.

b. Calculate the standard deviation of each to verify your rank ordering.

18. The following scores resulted from a biology exam:

<table>
<thead>
<tr>
<th>Scores</th>
<th>( f )</th>
<th>Scores</th>
<th>( f )</th>
</tr>
</thead>
<tbody>
<tr>
<td>95–99</td>
<td>3</td>
<td>65–69</td>
<td>7</td>
</tr>
<tr>
<td>90–94</td>
<td>3</td>
<td>60–64</td>
<td>6</td>
</tr>
<tr>
<td>85–89</td>
<td>5</td>
<td>55–59</td>
<td>5</td>
</tr>
<tr>
<td>80–84</td>
<td>6</td>
<td>50–54</td>
<td>3</td>
</tr>
<tr>
<td>75–79</td>
<td>6</td>
<td>45–49</td>
<td>2</td>
</tr>
<tr>
<td>70–74</td>
<td>8</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

a. What is the median for this exam?  
b. What is the mode?  

education