Central Tendency

Measures of Central Tendency

A measure of central tendency is a descriptive statistic that finds a single score that defines the center of a distribution:

- Most typical or representative score
- There are three common measures of central tendency:
  - the mode
  - the median
  - the mean

The Mean

The mean is:

- the arithmetic average of all the scores
  \( \frac{\Sigma X}{N} \)
- the number, \( \bar{X} \), that makes \( \Sigma (X - \bar{X}) \) equal to 0
- the number, \( \bar{X} \), that makes \( \Sigma (X - \bar{X})^2 \) a minimum
- the balance point of the distribution
- The mean of a population is represented by the Greek letter \( \mu \); the mean of a sample is represented by \( \bar{X} \)
Calculating the Mean

Calculate the mean of the following data:
1 5 4 3 2
Sum the scores (ΣX):
1 + 5 + 4 + 3 + 2 = 15
Divide the sum (ΣX = 15) by the number of scores (N = 5):
15 / 5 = 3
Mean = X = 3

The Mean

<table>
<thead>
<tr>
<th>X</th>
<th>X.X</th>
<th>X.X</th>
<th>X.X</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>5</td>
<td>25</td>
<td>25</td>
<td>25</td>
</tr>
<tr>
<td>4</td>
<td>16</td>
<td>16</td>
<td>16</td>
</tr>
<tr>
<td>3</td>
<td>9</td>
<td>9</td>
<td>9</td>
</tr>
<tr>
<td>2</td>
<td>4</td>
<td>4</td>
<td>4</td>
</tr>
<tr>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>2</td>
<td>4</td>
<td>4</td>
<td>4</td>
</tr>
</tbody>
</table>

Σ X = 55

The Mean

The sum of the distances from the mean = 0
-3 + -3 + -1 + 3 + 4 = 0
Characteristics of the Mean

- Changing the value of any score will change the mean
- Adding or removing a score (that is not the mean) will change the mean
- Adding or subtracting a constant from all scores will increase (addition) or decrease (subtraction) the mean by the constant
- Multiplying or dividing each score by a constant will change the mean in the same way

When To Use the Mean

- You should use the mean when
  - the data are interval or ratio scaled
  - Many people will use the mean with ordinally scaled data too
  - and the data are not skewed

The Median

- The median is another name for the 50th percentile
- It is the score in the middle; half of the scores are larger than the median and half of the scores are smaller than the median
How To Calculate the Median

- Conceptually, it is easy to calculate the median
  - There are many minor problems that can occur; it is best to let a computer do it
- Sort the data from highest to lowest
- Find the score in the middle
  - Index of middle score = \((N + 1) / 2\)
  - If \(N\), the number of scores, is even the median is the average of the middle two scores

Median Example

- What is the median of the following scores:
  
  10   8   14   15   7   3   3   8   12   10   9

- Sort the scores:
  
  15   14   12   10   10   9   8   8   7   3   3

- Determine the index of the middle score:
  
  index of middle = \((N + 1) / 2 = (11 + 1) / 2 = 6\)

- Middle score = median = 9

Median Example

- Sort the scores:
  
  3   3   7   8   8   9   10   10   12   14   15

- \(n = 11\)
- \(p = 50\)
- \(i = (n \times p) / 100 + 0.5 = (11 \times 50) / 100 + 0.5 = 6\)
- \(k = \text{integer part of } i = 6\)
- \(f = i - k = 6 - 6 = 0\)
- \(\text{Median} = (1 - f) \times \text{Obs}_k + f \times \text{Obs}_{k+1}\)
  
  \[= (1 - 0) \times 9 + 0 \times 10 = 9\]
When To Use the Median

- The median is often used when the distribution of scores is either positively or negatively skewed
- The few really large scores (positively skewed) or really small scores (negatively skewed) will not overly influence the median

The Mode

- The \textit{mode} is the score that occurs most frequently in a set of data

Bimodal Distributions

- When a distribution has two “modes,” it is called \textit{bimodal}
Multimodal Distributions

If a distribution has more than 2 “modes,” it is called multimodal.

When To Use the Mode

- The mode is not a very useful measure of central tendency.
  - It is insensitive to large changes in the data set.
    - That is, two data sets that are very different from each other can have the same mode.

When To Use the Mode

- The mode is primarily used with nominally scaled data.
  - It is the only measure of central tendency that is appropriate for nominally scaled data.
Relations Between the Measures of Central Tendency

- In symmetrical distributions, the median and mean are equal.
- For normal distributions, mean = median = mode.
- In positively skewed distributions, the mean tends to be greater than the median.
- In negatively skewed distributions, the mean tends to be smaller than the median.