CENTRAL TENDENCY

CHAPTER 3

YAKUP ARI,Ph.D.(C)
math.stat.yeditepe@gmail.com

OVERVIEW

- **REMEMBER!!!** The purpose of descriptive statistics is to summarize and organize a set of scores.

- One of methods of descriptive statistics is to find a single score that is the **average or representative** of the entire distribution. **This is known as central tendecy.**

OVERVIEW

- **The Central Tendecy** is a statistical measure to determine a single score that defines the center of a distribution.
- The goal of central tendecy is to find the single score that is to **most typical or representative of the entire group.**
- It is also useful for making comparisons between groups of individuals.
- There is **no single, standard procedure for determining central tendecy**.

OVERVIEW

- There are **three** different procedures for measuring central tendency. Namely,
  - 1) The Mean
  - 2) The Median
  - 3) The Mode

- **ATTENTION**!!! In order to decide which of three measures is best for any particular distribution, we should keep in mind that the general purpose of central tendency is to find the single most representative score.

THE MEAN

- The mean is generally known as the aritmetic mean.
- It is computed by adding all scores in a distribution and dividing by the number of scores.

- For a population, mean is identified by **μ**
- For a sample, mean is identified by ***M***

CHARACTERISTICS FOR THE MEAN

- Changing the value of any score will change the mean.

- If you add a new score or remove an existing score, the mean will generally change. However, adding or removing a score that is equal to the mean will not cause a change in the mean.

- If a constant value is added or subracted from each score, the same constant will be added or subracted from the mean.

- If every score in a distribution is multiplied or divided by a constant value, the mean will change in the same way.

MEDIAN

- Second measure of central tendency is the median.

  The **Median** is the score that divides a distribution **in half** so that 50% of the individuals in a distribution have scores at or below the median.

- For the median, there are no symbols or notation.

MEDIAN

- When n is an **odd** number, we calculate the median as follows:
  - List the scores from lowest to highest
  - The median is the middle score in the list
- When n is an **even** number, we calculate the median as follows:
  - List the scores from lowest to highest
  - The median is the average of the middle two scores.

MEDIAN

- When there are several scores with the same value in the middle of the distribution, we calculate the precise median for a continuous variable as follows:
  - Determine the real limits of the interval that contains the precise midpoint
  - Count the number of scores below the identified interval.
  - Find the number of additional scores needed to reach exactly 50%
  - Calculate a fraction=number of additional scores needed/ total number of scores in the interval
  - Add the fraction to the lower real limit of the interval

MODE

- The third measure of central tendency is the **mode.**

The **mode** is the score or category that has the greatest frequency

- It is a useful measure, it can be used to determine the typical or average value for any scale of measurement including nominal scale.
- Notice that it is possible to have two or more score that have the same highest frequency.

SELECTING A MEASURE OF CENTRAL TENDENCY

- The mean is usually the preferred measure of central tendency. Because,
  - it uses every score in the distribution, so it produces a good representative value
  - It is closely related to variance and standard deviation, the most common measures of variability. So this relationship makes the mean a valuable measure for the inferential statistics.
- However, there are specific situations in which it is impossible to compute a mean or in which the mean is not representative

WHEN TO USE THE MEDIAN

- There are four situations in which the median is a valuable alternative to mean.
  - **EXTREME SCORES OR SKEWED DISTRIBUTION:** Because the mean uses all of the scores in a distribution, it is affected by one or two extreme scores and the mean is displayed to the extreme scores.
  - **UNDETERMINED VALUES:** When there are undetermined scores in a distribution, it is impossible to compute $\sum X$. Therefore, we cannot compute the mean
  - **OPEN-ENDED DISTRIBUTIONS:** When the distribution is an open-ended distribution, it is impossible to compute $\sum X$. Therefore, we cannot compute the mean
  - **ORDINAL SCALE:** **Remember!!!** Ordinal scales allow us to determine the direction of difference but do not allow us to determine distance of difference. The mean is defined by the distance but median is defined by direction.

WHEN TO USE THE MODE

- There are three situations in which the mode is used to alternative to mean.
  - **NOMINAL SCALES**: Since the nominal scales do not measure quantity (distance or direction), it is impossible to compute mean or median.
  - **DISCRETE VARIABLES**: We can compute means. In this situaiton, the calculated means will be usually fractional values that cannot actually exist. Mode produced more sensible values.
  - **DESCRIBING SHAPE**: The mode gives an indication of the shape of the distribution. It is generally included as a suplementary measure along with the mean and the median.

CENTRAL TENDENCY AND THE SHAPE OF THE DISTRIBUTION

- **SYMMETRICAL DISTRIBUTIONS:**
  - For a perfectly symmetrical distribution, all three measures of central tendecy have the same value.
  - For a roughly symmetrical distribution, the three measures of central tendency are clustered together in the center of the distribution
- **SKEWED DISTRIBUTIONS:**
  - **For a positively skewed distribution**, the order of the three measures of central tendency from smallest to largest is as follows: MODE<MEDIAN<MEAN
  - **For a negatively skewed distribution,** , the order of the three measures of central tendency from smallest to largest is as follows: MEAN<MEDIAN<MODE