

Chapter Four Measurement and Descriptive Statistics

Measurement and Descriptive Statistics

- **Measurement**
- Measurement is the process of observing and recording the observations that are collected as part of a research effort.
- Statistically, it is the assignment of numerals(or symbols) to objects or events according to a set of rules.
- What rules?
- The rules you, the researcher, develop for the operational definition of your concept.

Measurement and Descriptive Statistics

- **Levels of measurement**
- A level of measurement is the precision by which a variable is measured. The four levels of measurement, from lowest to highest, are:
 - Nominal
 - Ordinal
 - Interval
 - Ratio

Measurement and Descriptive Statistics

- The **nominal level** of measurement describes variables that are categorical in nature. The characteristics of the data you're collecting fall into distinct categories.
- If there are a limited number of distinct categories (usually only two), then you're dealing with a discrete variable.
- If there are an unlimited or infinite number of distinct categories, then you're dealing with a continuous variable.
- Nominal variables include demographic characteristics like sex, race, and religion.
- ***The Nominal Scale***: consists of 'naming' observations or classifying them into various mutually exclusive categories. *Example*: Male – Female; Married – Unmarried; Well – Sick

Measurement and Descriptive Statistics

- The **ordinal level** of measurement describes variables that can be ordered or ranked in some order of importance.
- It describes most judgments about things, such as big or little, strong or weak.
- Most opinion and attitude scales or indexes in the social sciences are ordinal in nature.
- *Example:* Patients may be characterized as unimproved, improved, and much improved
- Individuals may be classified according to socioeconomic as low, medium and high.
- It is usually impossible to infer that the difference between members of one category and the next adjacent category is equal to the difference between members of that category and members of the category adjacent to it.

Measurement and Descriptive Statistics

- The **interval level** of measurement describes variables that have more or less equal intervals, or meaningful distances between their ranks.
- *Example:* Temperature: the unit of measurement is the degree and the point of comparison is the arbitrary chosen “zero degrees,” which doesn’t indicate a lack of heat.
- Also, if you were to ask somebody if they were first, second, or third generation immigrant, the assumption is that the distance, or number of years, between each generation is the same.
- All crime rates in criminal justice are interval level measures, as is any kind of rate.
-

Measurement and Descriptive Statistics

- The **ratio level of measurement** describes variables that have equal intervals and a fixed zero (or reference) point.
- It is possible to have zero income, zero education, and no involvement in crime, but rarely do we see ratio level variables in social science since it's almost impossible to have zero attitudes on things, although "not at all", "often", and "twice as often" might qualify as ratio level measurement.
- *Example:* height, weight, length, etc.

Measurement and Descriptive Statistics

- Advanced statistics require at least interval level measurement, so the researcher always strives for this level, accepting ordinal level (which is the most common) only when they have to.
- Variables should be conceptually and operationally defined with levels of measurement in mind since it's going to affect how well you can analyze your data later on.

Measurement and Descriptive Statistics

- **Descriptive statistics**
- Descriptive statistics fall into one of two categories: measures of central tendency (mean, median, and mode) or measures of dispersion (standard deviation and variance).
- Their purpose is to explore hunches that may have come up during the course of the research process, but most people compute them to look at the normality of their numbers.
- Examples include descriptive analysis of sex, age, race, social class, and so forth.

Measurement and Descriptive Statistics

- **Frequency Distributions**
- A frequency distribution gives a general picture of the value of a variable.
- It is often convenient, however, to summarize a quantitative variable still further by giving just two measurements, one indicating the average value and the other the spread of the values.
- The tendency of statistical data to get concentrated at certain values is called “Central Tendency” and the various methods of determining the actual value at which the data tend to concentrate are called measures of central tendency or average.

Measurement and Descriptive Statistics

- A descriptive measure computed from the data of a sample is called a *statistic*; a descriptive measure computed from the data of a population is called a *parameter*.
- *Average*: a figure that best represents the location of the distribution.
- There are three kinds of averages – each in a different way measuring the location of the distribution. These are the *Arithmetic Mean*, the *Median*, and the *Mode*.

Measurement and Descriptive Statistics

- **MEASURES OF CENTRAL TENDENCY**

- The most commonly used measure of central tendency is the mean.
- To compute the mean, you add up all the numbers and divide by how many numbers there are.
- It's not the average nor a halfway point, but a kind of center that balances high numbers with low numbers.
- For this reason, it's most often reported along with some simple measure of dispersion, such as the range, which is expressed as the lowest and highest number.
- The median is the number that falls in the middle of a range of numbers.
- It's not the average; it's the halfway point.
- There are always just as many numbers above the median as below it.

Measurement and Descriptive Statistics

- In cases where there is an even set of numbers, you average the two middle numbers. The median is best suited for data that are ordinal, or ranked. It is also useful when you have extremely low or high scores.
- The **mode** is the most frequently occurring number in a list of numbers.
- It's the closest thing to what people mean when they say something is average or typical.
- The mode doesn't even have to be a number.

Measurement and Descriptive Statistics

- It will be a category when the data are nominal or qualitative.
- The mode is useful when you have a highly skewed set of numbers, mostly low or mostly high.
- You can also have two modes (bimodal distribution) when one group of scores are mostly low and the other group is mostly high, with few in the middle.
- **MEASURES OF DISPERSION**
- In data analysis, the purpose of statistically computing a measure of dispersion is to discover the extent to which scores differ, cluster, or spread from around a measure of central tendency.
- The most commonly used measure of dispersion is the standard deviation.

Measurement and Descriptive Statistics

- You first compute the variance, which is calculated by subtracting the mean from each number, squaring it, and dividing the grand total (Sum of Squares) by how many numbers there are.
- The square root of the variance is the standard deviation.
- The standard deviation is important for many reasons.
- One reason is that, once you know the standard deviation, you can standardize by it.
- Standardization is the process of converting raw scores into what are called standard scores, which allow you to better compare groups of different sizes.
- Standardization isn't required for data analysis, but it becomes useful when you want to compare different subgroups in your sample, or between groups in different studies.

Measurement and Descriptive Statistics

- A standard score is called a z-score (not to be confused with a z-test), and is calculated by subtracting the mean from each and every number and dividing by the standard deviation.
- Once you have converted your data into standard scores, you can then use probability tables that exist for estimating the likelihood that a certain raw score will appear in the population.
- This is an example of using a descriptive statistic (standard deviation) for inferential purposes.