Variable

In statistics, a **variable** has two defining characteristics:

- A variable is an attribute that describes a person, place, thing, or idea.
- The value of the variable can "vary" from one entity to another.

Definition

A variable is a characteristic, often but not always quantitatively measured, containing two or more values or categories that can vary from person to person, object to object or from phenomenon to phenomenon.

Constant

- A logical opposite of a variable is a constant.
- A constant is a particular type of variable, which does not vary from one member of a group to another

Definition

The term constant refers to a property whereby the members of a group or category remain fixed and do not one from another.

Types of Variables

Variables can be classified as

- **Qualitative (or Categorical)**
- **Quantitative (or Numerical)**

- Qualitative variables take on values that are names or labels. Numerical measurement are not possible. The color of a ball (e.g., red, green, blue) or the breed of a dog (e.g., collie, shepherd, terrier) would be examples of qualitative or categorical variables.

- Quantitative variables are numeric. They represent a measurable quantity. For example, when we speak of the population of a city, we are talking about the number of people in the city - a measurable attribute of the city. Therefore, population would be a quantitative variable.

Level of Measurement

Level of measurement defines the amount of information contained in the data. According to measurement scale:

- Nominal
- Ordinal
- Interval Scales
- Ratio Scales

Nominal Variables

- Nominal variable – a categorical variable *without* an intrinsic (general) order
- Examples of nominal variables:
  - Where a person lives in the U.S. (Northeast, South, Midwest, etc.)
  - Gender (Male, Female)
  - Nationality (American, Mexican, French)
  - Race/ethnicity (African American, Hispanic, White, Asian American)

Ordinal Variables

- Ordinal variable—a categorical variable with some intrinsic order

- Examples of ordinal variables:
  - Agreement (strongly disagree, disagree, neutral, agree, strongly agree)
  - Rating (excellent, good, fair, poor)
  - Frequency (always, often, sometimes, never)

Interval Scales

- Interval data are measured and have constant, equal distances between values, but the zero point is arbitrary.
- There is no absolute zero.
- The zero isn't meaningful, it doesn't mean a true absence of something.

Example:

The difference between a temperature of 100 degrees and 90 degrees is the same difference as between 90 degrees and 80 degrees.

Ratio Scales

- A *ratio* variable, has all the properties of an interval variable, and also has a clear definition of 0.0.
- Ratio scales have an absolute zero
- When the variable equals 0.0, there is none of that variable.

Examples

Variables like height, weight, enzyme activity are ratio variables.

Interval Scales vs. Ratio Scales

- Temperature, expressed in F or C, is not a ratio variable. A temperature of 0.0 on either of those scales does not mean 'no heat'.

- However, temperature in Kelvin is a ratio variable, as 0.0 Kelvin really does mean 'no heat'.

- Another counter example is pH. It is not a ratio variable, as pH=0 just means 1 molar of H+. and the definition of molar is fairly arbitrary. A pH of 0.0 does not mean 'no acidity' (quite the opposite!).

- When working with ratio variables, but not interval variables, you can look at the ratio of two measurements.

- A weight of 4 grams is twice a weight of 2 grams, because weight is a ratio variable. A temperature of 100 degrees C is not twice as hot as 50 degrees C, because temperature C is not a ratio variable.

- A pH of 3 is not twice as acidic as a pH of 6, because pH is not a ratio variable.

Numerical Variables (Cont..)

Numeric variables have values that describe a measurable quantity as a number, like 'how many' or 'how much'.

A continuous variable is a numeric variable. Observations can take any value between a certain set of real numbers. The value given to an observation for a continuous variable can include values as small as the instrument of measurement allows. Examples of continuous variables include height, time, age, and temperature.

A discrete variable is a numeric variable. Observations can take a value based on a count from a set of distinct whole values. A discrete variable cannot take the value of a fraction between one value and the next closest value. Examples of discrete variables include the number of registered cars, number of business locations, and number of children in a family, all of which measured as whole units (i.e. 1, 2, 3 cars).

Univariate vs. Multivariate Data

Statistical data are often classified according to the number of variables being studied.

Univariate data

When we conduct a study that looks at only one variable, we say that we are working with univariate data. Suppose, for example, that we conducted a survey to estimate the average weight of high school students. Since we are only working with one variable (weight), we would be working with univariate data.

Multivariate data.

When we conduct a study that examines the relationship among more than two variables, we are working with multivariate data. Suppose we conducted a study to see if there were a relationship among the height, weight, and age of high school students. Since we are working with three variables (height , weight, age), we would be working with multivariate data.