AL-FARABI KAZAKH NATIONAL UNIVERSITY
FACULTY OF MEDICINE AND HEALTHCARE

Introduction to Biostatistics

Saktapov Akylbek
senior lecturer

Outline

- Definition of statistics
- Data and variables
- Type of variables

Definition of statistics

Statistics is a field of study concerned with the collection, organization, summarization, and analysis of data; and the drawing of inferences about a body of data when only a part of the data is observed.

When the data analyzed are derived from the biological sciences and medicine, we use the term **biostatistics** to distinguish this particular application of statistical tools and concepts.

Training in statistics has been recognized as “indispensable” for students of medical science. For example:

- if we want to establish cause and effect relationship, we need statistics.
- if we want to measure state of health and also burden of disease in community, we need statistics.

Statistics are widely used in epidemiology,

- clinical trial of drug vaccine program planning community medicine
- health management
- health information system etc..

The knowledge of medical statistics enables one to develop a self- confidence & this will enable us to become a good clinician, good medical research worker, knowledgeable in statistical thinking.

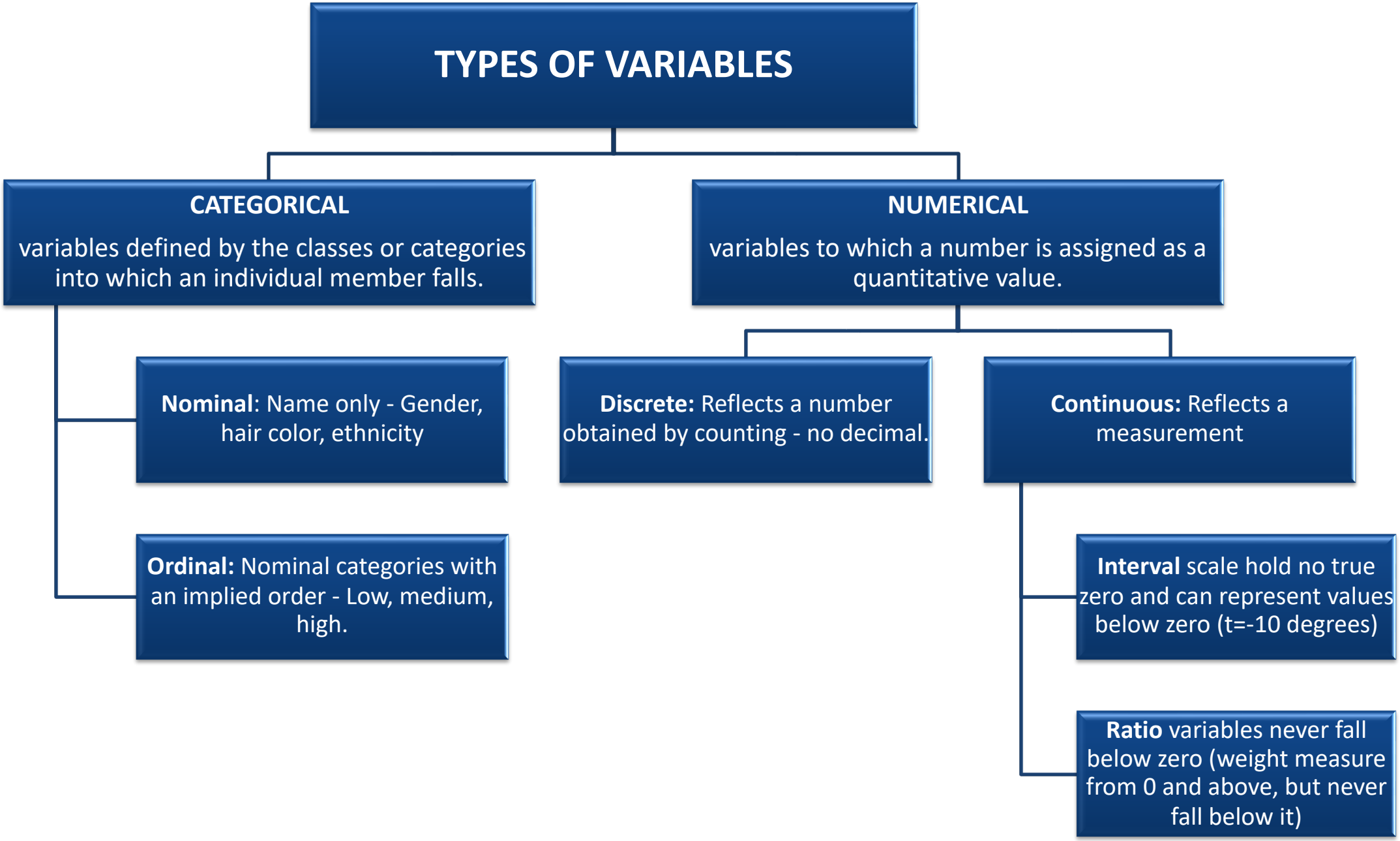
Data and Variables

The raw material of statistics is **data**. For our purposes we may define data as numbers.

Data are numbers, numbers contain information, and the purpose of statistics is to investigate and evaluate the nature and meaning of this information.

Data are often discussed in terms of variables, where a variable is: any characteristic that varies from one member of a population to another.

A simple example is height in centimeters, which varies from person to person.


TYPES OF VARIABLES
CATEGORICAL
variables defined by the classes or categories into which an individual member falls.
NUMERICAL
variables to which a number is assigned as a quantitative value.
Nominal: Name only - Gender, hair color, ethnicity
Ordinal: Nominal categories with an implied order - Low, medium, high.
Discrete: Reflects a number obtained by counting - no decimal.
Continuous: Reflects a measurement
Interval scale hold no true zero and can represent values below zero (t=-10 degrees)
Ratio variables never fall below zero (weight measure from 0 and above, but never fall below it)

Categorical or Numerical?

- Score on a placement exam
- Preferred restaurant
- Dollar amount of a loan
- Height
- Level of satisfaction
- Salary
- Number of applicants
- Ethnic origin

- Numerical
- Categorical
- Numerical
- Numerical
- Categorical
- Numerical
- Numerical
- Categorical

Independent vs. Dependent Variables

- The independent variable is what the researcher studies to see its relationship or effects.
  - Presumed or possible cause
- The dependent variable is what is being influenced or affected by the independent variable
  - Presumed results
- Independent variables may be either *manipulated* or *selected*
  - A manipulated variable is a changed condition the researcher creates during a study, also known as an experimental or treatment variable
  - A selected variable is an independent variable that already exists

INDEPENDENT VARIABLE

What I CHANGE

DEPENDENT VARIABLE

What I OBSERVE

The Relationship between Independent and Dependent Variables

Example

The temperature of water was measured at different depths of a pond.

- Independent variable?
  depth of the water

- Dependent variable?
  temperature

Example

If the temperature of water is higher, then an egg will boil faster.

- Independent variable?
  temperature of water

- Dependent variable?
  time to cook an egg

Thank you for your attention!!!