

INTRODUCTION TO STATISTICS

Tudor Călinici

2026

Objectives

- ▣ To differentiate between different types of statistical applications
- ▣ Know the specific terms in statistics
 - ▣ Variable
 - ▣ Population
 - ▣ Sample
 - ▣ Size
 - ▣ Bias
 - ▣ Representative sample
 - ▣ Statistical series

Statistics

- ▣ Scientific area allowing the study of the phenomena which are characterized mainly by the variability

The variable

- ▣ The variable is the characteristic which may take different values from object to object (or at the same object in time)

- ▣ The opposite of the variable is the constant

$$\Pi=3.14$$

Statistics

- ▣ Descriptive statistics – collecting data, presenting data, computing numerical characteristics of the data in order to obtain information about the groups
- ▣ Inferential statistics – finding information about the population starting from a fraction of it

Statistics

- ▣ Is about statistical sets not about individuals

Statistical sets

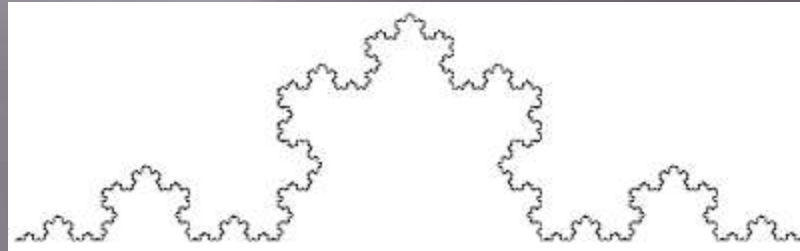
- ▣ Population
 - ▣ All the elements with a specific characteristic
 - ▣ Population size: Number of objects in population
- ▣ Sample
 - ▣ Sample– part of the population
 - ▣ Sample size: Number of objects in sample

Population vs Sample

- ▣ In practice, in statistical studies we'll use samples
- ▣ Why:
 - The population size may be infinite.
 - Sometimes we can obtain information about only part of the population.
 - Economic reasons.
 - Limited resources.

Representative sample

- ▣ A good sample should represent, as accurately and faithfully as possible, the entire population in which a particular characteristic is to be studied.



- ▣ Is the good sample representative? Is the representative sample good?
- ▣ Representative sample
 - Quantitative condition
 - Qualitative condition

Methods of sampling

- ▣ Probabilistic approach
 - Randomize
 - Systematical sampling
- ▣ Stratifying
- ▣ Clustering
- ▣ Non - Probabilistic approach
 - Convenience
 - Snowball
- ▣ BIAS

The variable

- ▣ The variable is the characteristic which may take different values from object to object (or at the same object in time)
- ▣ Variable – Name and the set of values

Example

- ▣ Characteristics of interest (Variables) :
 - Gender; Marital status

- ▣ Set of values:
 - Gender = {Female, Male}
 - Marital status= {Not married, Married, Divorced, Widow / Widower}



Variables

- ▣ Quantitative variables – can be measured
- ▣ Qualitative variables – can not be measured

Quantitative variables

- ▣ Continuous
 - Weight, height, Systolic Blood Pressure, Diastolic Blood Pressure, ...
- ▣ Discrete
 - Number of days in hospital, Quantity of the drug prescribed, Number of children...
- ▣ We can do math using those variables

Qualitative variables

- ▣ Nominals
 - Name, nationality, hair color, type of blood, ...
- ▣ Ordinals
 - Health status – poor, better, good, very good
 - Education – kindergarten, primary school, middle school, high school, bachelor's degree, master's degree, doctorate
 - ...
- ▣ Dichotomous (binary)
 - Gender (male/female) Health Status (Healthy / Sick) ..
- ▣ We can not do math with the values of the qualitative variables

Scales

- ▣ Qualitative variables
 - Nominal
 - Ordinal
- ▣ Quantitative variables
 - Interval
 - Ratio

Changing the variable type

- ▣ We can transform quantitative variables in qualitative variables
- ▣ Example: From the quantitative variable “age,” with values ranging from 0 to 120 years, we can derive the variable “age group,” with categories such as newborn, infant, preschooler, school-age child, and so on.

Statistical series

- ▣ A statistical series is a sequence of observations of one or more variables, recorded for the statistical units in a sample.

A screenshot of a Microsoft Excel spreadsheet. The spreadsheet contains a table with 12 rows of data. The columns are labeled: ID, Nume, DataN, Profesia, Orasul, Stciv, and Salariul. The data represents a statistical series of observations for 12 individuals. The spreadsheet is titled 'DrExcel' and shows the standard Excel interface with a ribbon at the top.

ID	Nume	DataN	Profesia	Orasul	Stciv	Salariul
1	Ionescu	14.12.70	inginer	Ploiesti	c	2200
2	Vasilescu	05.09.60	medic	Bucuresti	n	1600
3	Radulescu	02.03.80	manager	Cluj	d	3500
4	Antonescu	08.07.50	asistent	Iasi	v	900
5	Marinescu	01.08.75	economist	Brasov	c	1900
6	Ionescu	14.12.70	inginer	Ploiesti	c	2200
7	Vasilescu	05.09.60	medic	Bucuresti	n	1600
8	Radulescu	02.03.80	manager	Cluj	d	3500
9	Antonescu	08.07.50	asistent	Iasi	v	900
10	Marinescu	01.08.75	economist	Brasov	c	1900
11	Ionescu	14.12.70	inginer	Ploiesti	c	2200
12	Vasilescu	05.09.60	medic	Bucuresti	n	1600

IMPORTANT – STATISTICS

- ▣ Collects and presents data.
- ▣ Selects and calculates the appropriate parameters.
- ▣ Selects and performs the appropriate statistical tests and interprets the results.

IMPORTANT – STATISTICS

- ▣ Does not define or select the population/sample.
- ▣ Does not select the appropriate variables for the situation.
- ▣ Does not measure accurately (most of the time).
- ▣ Does not formulate medical hypotheses.

Statistics – First year medical students

- ▣ Understanding statistical determinism in medical literature.
- ▣ Distinguishing between medical truth and statistical results.
- ▣ Performing simple statistical analyses – calculating parameters and interpreting results.
- ▣ Reading medical texts containing advanced statistical concepts with the help of references.
- ▣ Understanding and correctly using statistical terminology