THE STATE UNIVERSITY OF MEDICINE AND PHARMACY "NICOLAE TESTEMITANU“

DEPARTMENT of
MANAGEMENT AND PSYCHOLOGY

Introduction to Biostatistics. Nominal and Ordinal Data Summarizing.

Nina Globa,

Master in Public Health Management, lecturer

Discussion topics

1. Statistics and Biostatistics – interpretation of the concepts.
2. Reasons for learning biostatistics
3. The history and stages in the evolution of statistics
4. Theoretical foundations of biostatistics, objectives and compartments
5. Methods of investigation in biostatistics
6. The basic notions in biostatistics: Statistical totality, unit of observation, statistical signs
7. Scales of measurement.
8. Nominal and ordinal data summarizing. The Relative indicators: rates; ratios; proportions

1. Statistics

- ***Statistics*** **is a universal science that deals with the study of the quantity of all phenomena from society and nature in a correlation with their qualitative part .**

Mining of word *statistics*: data or numbers; the process of analyzing the data; the description of a field of study.

It derives from the Latin word *status* , meaning "*manner of standing* " or *"position"*

- Working with statistics involves using statistical methods that summarize data and using statistical procedures to reach certain conclusions that can be applied to patient care or public health planning.

Areas of statistics

The application of statistics includes:

- Business, marketing, economics, agriculture, education, psychology, sociology, anthropology, biology, medicine and other health care disciplines.
- We use the terms *Biostatistics and Biometrics* to refer to the application of statistics in a heath –related fields

Biostatistics

- ***Biostatistics*** is a social science that studies the quantitative parts of health phenomena in correlation with their qualitative parts in concrete conditions of space and time

Biostatistics has two important features:

1. Study of mass phenomena, not in particular
2. Study of phenomena in concrete conditions of space and time

2. Reasons for learning biostatistics

- **Evaluating the Literature** (reading the literature begins early in the training of health professionals an continues throughout their careers. They must understand biostatistics to decide whether they can believe the results presented in the literature. Exe. Williamson and colleagues (1992) analyzed 4235 published medical study and determined that only about 20% of them met the assessors' criteria for validity.
- **Applying study results to patients care.** This is the major reason of clinicians for reading the medical literature. They want to know which diagnostic procedures are the best, which methods of treatment are optimal.

Reasons for learning biostatistics

- **Interpreting Vital Statistics.** Physicians must be able to interpret VS in order to diagnose and treat patients effectively. Vital events are births and deaths.

- **Understanding Epidemiologic Problems.** This information helps practitioners to make diagnoses and develop manager plans for patients. Epidemiologic data reveal the prevalence of a disease, its variation by season and by geographic location, and its relation to certain risk factors.

Reasons for learning biostatistics

- **Interpreting information about new drugs and equipment.**
- **Identifying the correct diagnostic procedures to use** – is a necessity in making decisions about patient care, based on sensitivity and specificity.
- **Being informed.**
- **Appraising Guidelines.**
- **Evaluating study protocols and articles.**
- **Participating in research projects** (the license thesis project, resident research projects, other scientific projects) .

3. The history and stages in the evolution of statistics

- Stage of practical statistics
- Stage of descriptive statistics
- Stage of political arithmetic
- Stage of probabilistic statistics
- Stage of modern statistics

Stage of practical statistics

- During Sumerian, Egyptian, Greek and Roman civilization, data were obtained for the primary purpose of taxation and enrollment in the army. The records show incipient land, number and natural population movement.
- In the Roman Empire were made regular population records, customs records are prepared and cadastral records.
- In the Middle Ages, religious institutions often hold records of births, deaths and marriages.
- Statistical evidence, even if it amounts of simple records of facts, provides data to inform state regarding tax, military and administrative issues.

Stage of descriptive statistics

- Is a transition from "simple counting of facts" to "description of the facts".
- Representatives of the "state description" current, Hermann Conring, Martin Schmeitzel, Gottfried Achenwall introduced for the first time the term of statistics, developed methods of investigation of social and economic phenomena.

Stage of political arithmetic

- It belongs to the period when in England has appeared a particular statistic, known as the "political arithmetic".
- It deals with the analysis, mathematically processing of the collected data and formulating of regularities or forecasts.
- John Graunt identified regularities of population and demographic phenomena.
- William Petty (the father of modern political economy) used quantitative methods to study the social and economic phenomena.
- Edmund Halley tried to estimate the population, developed the first mortality table and introduced the concept of probable duration of life.

Stage of probabilistic statistics

- Name of *political arithmetic* was substituted with *statistics* and acquired a new dimension by introducing probabilistic calculations.
- It was formulated the law of large numbers.
- Among prominent representatives of this phase are: J. Bernoulli, PS Laplace, K.F. Gauss, S.D. Poisson, A. Ciuprov etc.

Stage of modern statistics

- It appeared in the end of XIX century.
- In this stage were created national and international statistical offices, organized international statistical conferences, journals and introduction of statistics in secondary and university education.
- At this stage F. Galton, K. Pearson, M. G. Kendall, G.U.Yule, C.E. Spearmen, R.A. Fisher etc., underpinned theory and practice of statistical correlation, factorial analysis and statistical experiments.
- They discussed in particular the problems of distribution, specification and estimation.

4. Theoretical foundations of biostatistics

- Dialectics of matter
- The general theory of statistics
- Economic laws
- Bases of medicine and its branches

Objectives of biostatistics

1. Studying of health status of population

- Reproduction of population
- Morbidity and disability
- Parameters of physical development

Objectives of biostatistics

2. Studying the link between the environment and the influence of risk factors on health.
3. Collection and analysis of data about institutions and health professionals.
4. Evaluation of the effectiveness of prevention and treatment methods.
5. Analysis of planning, economy and financing of health care.

Compartments of biostatistics

- Health Statistics
- Healthcare Statistics
- Statistics of prevention, treatment and other activities in medicine

5. Methods of investigation in biostatistics

- Historical
- Mathematical
- Statistical
- Epidemiological
- Sociological
- Psychological
- Experimental
- Economic etc.

Data collection methods

- Census of population
- Monographs (books with scientific results of somebody research)
- Sociological methods: questionnaires, interviews
- Extracting information from an official statistical form (clinical observation sheet, ambulatory sheet etc.)
- Direct observational method (objective examination, percussion, palpation, auscultation)
- Investigational method (X-ray examination, level of hemoglobin and leucocytes etc.)

6. The Basic notions in biostatistics

- Statistical community/totality
- Unit of observation
- Statistical signs, variables
- Statistical data
- Statistical indicators

Statistical community

- A group that includes a great number of things, objects, called elements (units of observation), relatively homogeneous, taken together on the basis of common factor in the boundaries of space and time.

Types of statistical community. There are:

- General (integral) statistical community (N)
- Partial statistical community (sample, n)

Statistical totalities

- General (integral) Statistical community (N) is the summarizing of all observation units, according to the purpose of the study, which are possible in specific boundaries of space and time.
- Partial Statistical community (sample, n) includes a representative part of integral totality

The requirements for the PSC are:

- To have available the basic characteristic features of the integral one;
- To have a sufficient number of observations, n.

Unit of observation

- Is the primary element of statistical community
- The number of unit of observation determines the volume of sample.
- Observation unit's characteristics are called signs(characteristics or variables).

Classification of variables (characteristics, signs)

- Depending on characters or expressions can be:
  - **Attributive (qualitative, descriptive)**
  - **Numerical (quantitative)**
- Depending on type of influence:
  - **Factorial** variables – **are variables under whose influences other variables are changed (age, gender etc.)**
  - **Result** variables – **reproduce those results and are changed under the influence of factorial** variables (exe. disease etc.)

Classification of variables

- Quantitative characteristics can be:
  - Continuous – take any value (even decimal)
  - Discontinuous – take the appearance of fixed values, usually of integer numbers (ex. Number of the children in a family)
- Qualitative characteristics may be:
  - ✓Nominal – with a natural ordination of categories (ex. gender: male & female)
  - ✓Ordinal – with a simple ordination of categories (ex. Level of education, results of treatment)

7. Scales of measurement

- The scale for measuring a characteristic has implications for the way is displayed and summarized
- The scales of measurement that occur most often in medicine are:

1. Nominal
2. Ordinal
3. Numerical

Nominal scales

Are used for the simplest level of measurement when data values fit into categories.

Could be:

- dichotomous or binary (yes/no; male/female)
- categorially as classification of some things, as for ex. types of diseases

Ordinal scales

- Are similar to nominal but describe stages or degrees of severity (level of education, results of treatment, the stages of tumors)

Numerical Scales

- Quantitative observations

The are two types of numerical scales:

1. Continuous – has values on a continuum (age)
2. Discrete – has values equal to integers (number of children, of pregnancies, of operations etc.)

8. Nominal and ordinal data summarizing.

- The statistical data obtained during the research express individual or group proprieties that usually are presented using absolute values, which are difficult to assess, because there are missing comparison elements and are used for analyzing very rarely phenomena.
- for more detailed analyzes are used relative values, indicators of central tendency , indicators of variation, hypothesis testing and other indicators.

Types of statistical indicators

I. Primary (Absolute)– which are measured in the unit of measure of the characteristics (kg, cm, km).

II. Secondary

- Relative: rates; ratios; proportions, especially for nominal and ordinal scales but also for numerical.
- The measures of central tendency: arithmetic mean (average); mode; median for numerical scales.

Relative indicators

All the relative indicators are based on the same formula: $(x/y) \times 10^n$,

where n could be:

1 – (the product =10)

2 – (the product = 100); the unit of measure is percent , the symbol – %;

3 – (the product = 1000); the unit of measure promilles, per thousand; the symbol – ‰;

4 – (the product = 10000); the unit of measure is prodecimilles, per ten thousands; the symbol – ‱;

5 – (the product = 100000), the unit of measure is procentimilles , the symbol – ‰0;

Relative indicators: *Ratios*

- Ratio is used to compare the occurrence of a variable in two different groups.
- Usually there are two completely independent groups. If one of this group is included in the other as a part, this ratio is called a proportion .
- It is expressed in numbers or coefficients (ex. female / male is 14 to 10; (1/1; 1,2/1), perdecimile, percentimilles.

Relative Indicators : *Proportions*

- Determine the quantitative distribution of the general totality , of components to their total – the structure.
- A type of ratio, in which *x* is part of *y*.
- Is usually expressed in percentages – %

Relative Indicators : *Rates*

- Rates measure the occurrence of an event in a population over time
- The basic formula for a rate is

$$\textit{Rate} = \frac{\text{Number of cases occurring during a given time period } x10^{n}}{\text{Population at risk during the same period}}$$

Rates

- Are always specific to a particular population. They reflect grouping of people based on time, place and person.
- Time: year, month, week, day or hour;
- Place: country, state, city, school, institution;
- Person: age, sex, profession etc.
- The units of measure are usually ‰, ‰‰, ‰‰‰‰
  - ***Ratios, proportions and rates*** are used to describe morbidity and demography.

- Thanks for your attention!!!