Biostatistic

Lecture - 1

- It is a fact that we are living in the information age (information revolution) for example every year about 0.5 million new articles published only in the medical field yearly, we are bombarded by data in our everyday lives, thus we need to know how to obtain these information, how to analyze, and how to interpret, these information (which is called data). Data are available in the form of numbers (values).

- **Statistics:**
- The term statistics has several meanings: data or numbers, the process of analyzing the data, and the description of a field of study. Statistics also known as the science of collecting, organizing, and interpreting numerical facts. It derives from the Latin word status, meaning manner of standing" or "position".
- Statistics is that field of science concerned with the collection, organization, presentation, and summarization of data, and drawing of inferences about whole body of data when only a small part of the data is observed or examined or considered.
- Historically; the earliest origin of statistics lies in the desire of rulers to count the number of inhabitants or measure the value of taxable lands in their domains. The new discipline of statistics took shape in the twentieth century to be *the science of data.*

- *Biostatistics:*

  It is that field of statistics in which the data being analyzed were derived from the biological sciences and medicine. There are two main objectives from statistics;

1. *In, which we are, concerned with only collection, organization, presentation and summarization of data that is called descriptive statistics.*
2. *In which the objective is to reach a decision about a large group of data by examining only a small past of the data, and it is called inferential statistics (analytic statistics).*

- *Data:*

  The raw material of statistics is called data; it is obtained either as measurement or as process of counting. It represents a collection of values (single of data is called datum).
- *Value: It is the numerical representative of the measurement of the variable.*
- *Sources of data:*

  The need for statistical activities is motivated by the need to answer a question, that need an appropriate approach, and the search for and the search for suitable data to serve as the raw material for that investigation, such data are usually available in the form of one or more of the following sources;

1. Routinely kept records.
2. Experiments.
3. External sources, in form of published reports, commercially available data banks, or the research literature.

- **Variable:**
- Any characteristic that can take different values in different occasions, places, persons, and time. It is labeled as **variable** e.g. Height, weight, age, etc...
- Variables are one of two types

1) **Quantitative** variable (numerical); is that variable that can be measured by units such as height, weight, age, etc…
2) **Qualitative** variable (categorical); is that variable that cannot be measured by usual sense or units, it can only be assessed by number or percentage e.g. Sex, ethnic group, color of the eye, race, education, occupation, type of disease.

- **Quantitative variables are of two main types:**

1) **Discrete** quantitative variable; characterized by gaps or interruptions in the values that it can assume, these gaps or interruptions indicate the absence of values between particular values that the variable can assume, e.g. Daily admission of patients to hospital,.
2) **Continuous** quantitative variable; it is also called continuous random variable, it does not posses the gaps or interruptions characteristic, it has fractions of units, and the variable can assume any value within a specified interval such as height.

- **Measurements and measurement scales:**
- These rules include;

1) **Nominal** scale (male-female, well-sick, under 65 years- 65 and above, child-adult, and married-unmarried). Most of the nominal data in the form of **binary** or **dichotomous** "the response is one of two…Yes or No".
2) **Ordinal** scale (high-intermediate-low, not smoker, light, moderate, heavy smoker, Social class I,II,III,IV&V) or it is as **rank**-ordered scale.
3) **Interval** scale (Age as 20-, 30-, 40-, 50-)
4) **Ratio** scale (determine the quality of ratio or interval)

- **Population:**

It is the largest collection of entities of which we have an interest at a particular time, sharing at least one characteristic in common. Populations may be **finite** or **infinite**.

- **Sample:**

The sample may be defined as a part of population, subset of population chosen in a representative way to be as much as possible representative for the population (random, or non random). The method applied to collect a sample is called sampling.

- **Uses of statistics:**

1) To measure the health state of the community and identify its health problems.
2) To compare health condition (status) of a community with others.
3) For planning of health services.
4) For evaluation of health services.
5) For estimating the future needs.
6) For research.
7) Evaluating the literature
8) Applying study results to patient care
9) Interpreting vital statistics
10) Understanding epidemiological problems
11) Interpreting information about drugs and equipment
12) Using diagnostic procedures
13) Being informed
14) Appraising guidelines
15) Evaluating study protocols and articles
16) Participating in or directing research projects

END