



Biostatistics

1. General concept of Statistics

1. Introduction to Biostatistics:

1.1 Definitions:

1. Statistics: is a branch of mathematics dealing with data collection, organization, analysis, interpretation and presentation.

2. Biostatistics: when the analyzed data are derived from the biological sciences and medicine, we use the term biostatistics, especially in medicine and pharmacy, for example accidents statistics, medical and biological statistics, etc. The collection, summarization, and analysis of data from those experiments and inference from the results. A major branch is medical biostatistics, which is exclusively concerned with medicine and health.

3. Observations(Data): Any recorded information about certain phenomena whether it is numerical or Categorical from which conclusions may be drawn, called observations. Such as blood pressure, heart rate and the weights of preschool children.

Sources of data

1. *Routinely kept records*
2. *Surveys*
3. *Experiments*

4. Variable: *An observed characteristic that takes different values in different subjects. If the values of the variable cannot be exactly predicted for the object. Then we call it random variable.*

There are two types of **variables** :

- 1) **Qualitative:** *is that phenomena whose observations can not be measured by numerical numbers, but can be categorized only.*

For example,

- *Tossing a coin: $x = \text{head, tail}$*
- *Gender: $x = \text{male; female}$*

- 2) **Quantitative:** *is that phenomena whose observations can be measured by numerical numbers.*

For example,

- *Casting a die : $x = 0, 1, 2, 3, 4, 5, 6$*
- *exam grades : $x = 0, 1, \dots, 100$*

many phenomena can be measured by numerical numbers such as weight, height , time,...



The Quantitative data may be Discrete or Continuous.

i) **Discrete data** : When it is defined on countable set of data whether it is finite or infinite.

for example

- Casting a die : $x=0,1,2,3,4,5,6$ (countable and finite)
- Natural number : $x= 1,2,3,4,...$ (countable and infinite)

ii) **Continuous data**: When it is defined on uncountable set of data.

for example

- Selecting a real number from interval $(0,1)$: $0 < x < 1$
- The heights of students in cm at certain university may be : $1.40 \leq x \leq 1.90$

5. Measurement of Variables:

A measurement variable is a variable that can be measured and given a number, like 4 mm, 3 yards or 42. Measurement variables are sometimes called numeric variables.

There are four different levels of measurement:

1- **Nominal level of measurement**: classifies data into mutually exclusive exhausting categories in which no order or ranking can be imposed on the data.

Examples:

- Blood Type (A,B,AB,O).
- Suppose there are data about people belonging to three different gender categories. In this case, the person belonging to the female gender could be classified as F, the person belonging to the male gender could be classified as M, and transgendered classified as T.

2- **Ordinal level of measurement**: classifies data into categories that can be ranked, however, precise differences between the ranks does not exist.

Examples:

- Feeling (Very Unhappy, Unhappy, Ok, Happy , Very happy),
- Suppose a student scores the highest grade of 100 in the class. In this case, he would be assigned the first rank. Then, another classmate scores the second highest grade of a 92; she would be assigned the second rank. A third student scores an 81 and he would be assigned the third rank, and so on. The ordinal level of measurement indicates an ordering of the measurements.

3- **Interval level of measurement**: ranks data and precise differences between units of measure do exist, however, no meaningful zero. Moreover, ratios can not be computed

Examples :

- Celsius temperature
- The date.



- an interval level of measurement could be the measurement of worry in a student between the score of 10 and 11, this interval is the same as that of a student who scores between 40 and 41.

4- Ratio level of measurement: poses all characteristics of Interval measurement, and there is a true zero. These variables can be meaningfully added, subtracted, multiplied, divided (ratios)

Examples:

- Height
- Weight

4. **Population:** Population: is a set of similar items or events which is of interest for some question or experiment. A population can be observations from animals, machines, places, or cells.

An the past years the word "Population" referred to observations obtained on statistical studies involving people. Today population refer to observation relevant to anything interest whether it be group of people, animals, or object sample.

For example, we are interested in the weights of all the children enrolled in a certain county elementary school system, our population consists of all these weights.

❖ Populations may be finite or infinite.

- If a population of values consists of a fixed number of these values, the population is said to be finite.

For example numbers of patients you examine daily.

- A population consists of an endless succession of values. the population is an infinite one.

For example the social population.

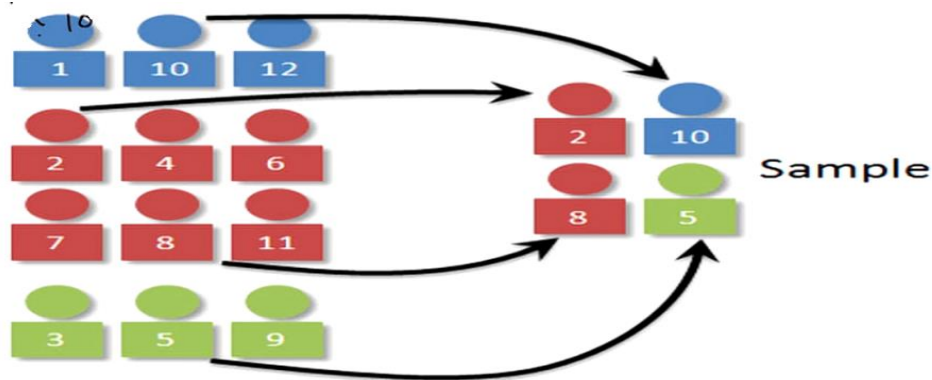
7. **Sample:** consists one or more observations drawn from the population.

That is a sample is a subset of a population. On practice, examination a whole population may be impossible. expensive and need effort, time, and money. Therefore the statistician depend on a subset of observations (Sample) from the population under consideration. The study of a Sample properties to make inference or conclusion about the population .

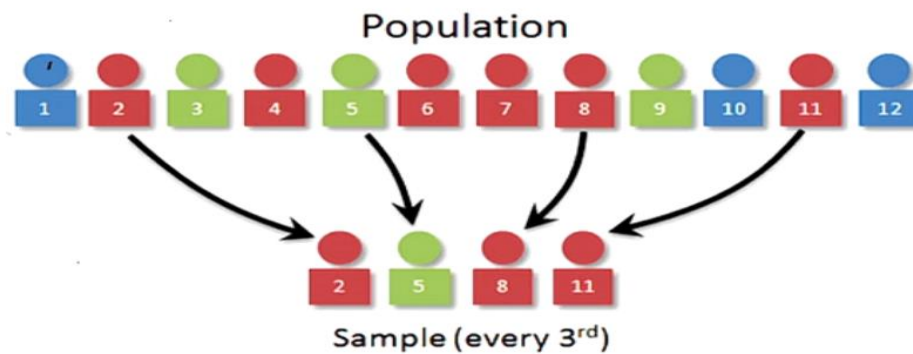
In Example to determine the average life of a certain light- bulb produced a certain factory, it would be impossible to test all bulbs ,so economic advantages and quickly information to have if we study average life of such bulbs by testing a sample.

Sampling Methods

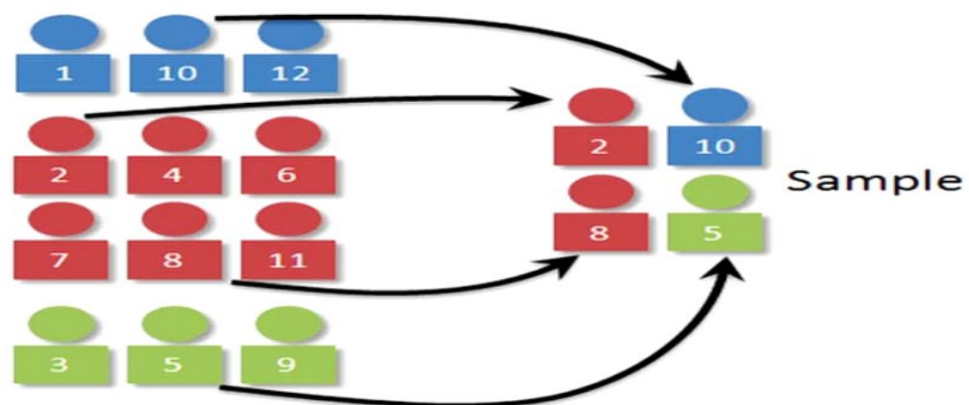
- 1- **Random Sampling:** random samples are selected by using chance methods or random numbers. Every subject in the population has the same chance of being selected.



- 2- **Systematic Sampling:** can be obtained by numbering each subject of the population and selecting every k th subject.



- 3- **Stratified Sampling:** can be obtained by dividing the population into groups (strata) according to some characteristic that is important for the study, then sampling from each group.



- 4- **Cluster Sampling:** here the population is divided into groups (clusters) by some means such as geographical areas. Then the researcher randomly selects some of these clusters and uses all members of the selected clusters.

