



Introduction To Biostatistics

By

Dr. Lamyia Al-Azzawi PhD. Community Dentistry

Department of Pediatric, Orthodontic and Preventive Dentistry (P.O.P)

Biostatistics.

It is a combination word made from biology and statistics)

The application of statistics to a wide range of topics in biology problems, including:

- ❖ Public health.
- ❖ Medicine.
- ❖ Ecological and environmental.

Definition

It is the science that provide a way of processing number or analyzing the data that have been collected . The purpose of statistics is to make an implication or an assumption about population which deals with development and application of the most appropriate methods for the:

- ❖ Collection of data.
- ❖ Presentation of the collected data.
- ❖ Analysis and interpretation(understanding) of the results.
- ❖ Making decisions on the basis of such analysis .

Types of Statistics.

1.Descriptive statistics are used to describe and summarize data. Their objective is to communicate results, without generalizing beyond the sample ,to any population .

Some ways in which results are communicated are through

- a. measures of central tendency (mean, median ,and mode)
- b. measures of dispersion

2. Inferential statistics are used to apply information from the sample to a larger population.

Challenge



Much of life is composed of a systematic component (i.e., signal) and a random component (i.e., error or noise)

- Example:
 - Smoking is associated with lung cancer.
 - Yet not everyone that smokes, gets lung cancer, and not everyone that gets lung cancer, smokes
 - Yet we know that there is an association (a systematic component)



Sources of Data

We search for **suitable data** to serve as the **raw material** for our investigation.

Such data are available from one or more of the following **sources**:

1- Routinely kept records.

For example:

- **Hospital** medical records contain huge amounts of information on **patients**.
- **Hospital** accounting records contain a wealth of data on the **facility's business**
- **activities**.

2.External sources.

The data needed to answer a question may already exist in the form of published reports, commercially available data banks, or the research literature, i.e. someone else has already asked the same question.

3- Surveys:

The source may be a survey, if the data needed is about answering certain questions which is known as questionnaire is prepared this method is easy to adopt when a large wide geographic area need to be covered. It is relatively cheap and fast usually the question should be short and easy to understand. There should be no uncertainty while answering the questions as far as possible questions should be close ended.

.

4- Experiments.Frequently the data needed to answer a question are available only as the result of an experiment.

5.Oral health examination

Collecting data regarding oral health diseases in this method oral examination should be done to get the information ,and it is usually conducted by the dentists , veterinarian ,technicians ,the investigators.

Populations and Samples

Studying populations is too expensive and time-consuming, and thus impractical.

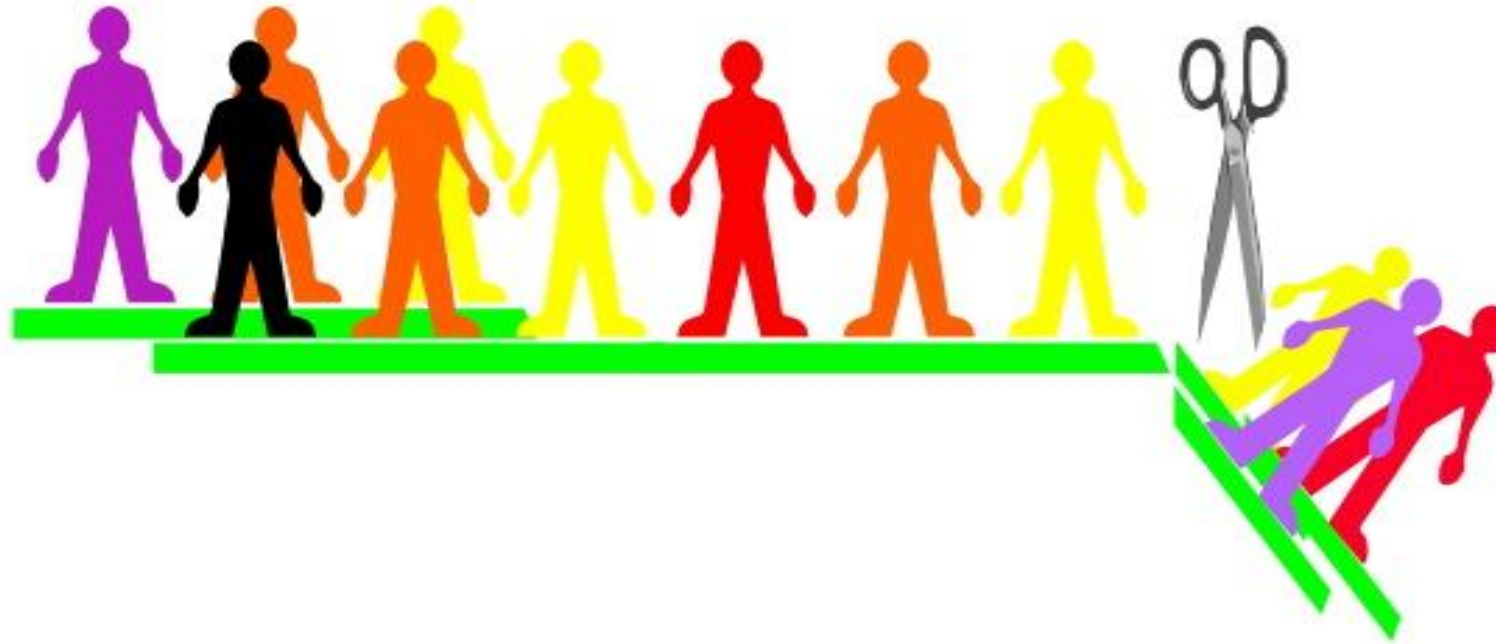
A sample is a portion of the population that, if properly selected can provide meaningful information about the entire population.

And thus by looking at the characteristics of the sample (statistics), we may learn something about the characteristics of the population (parameters)

*IF THE POPULATION IS
HOMOGENEOUS*



*IF THE POPULATION IS
HETEROGENEOUS*



The advantages of sampling over the census enumeration

Census

1. Information is collected from all the individuals in the population
2. Cost of organization and execution will be more
3. Require more time and personnel for collection and analysis .
4. Lesser accuracy and completeness

Sampling

1. Information is collected from the units in the sample.
2. Cost is less since sample is smaller than the population.
3. Requires lesser time and personnel for collection and analysis .
4. More accuracy and completeness

Types of Samples

1. Purposive (Judgmental) selection

This type of sample provide a sample through personal judgment ,of subjects who would be most representative of population, it is easy to carry out and does not need the preparation of sampling frame.

For example:

Studying oral cancer in men forty representative patients may be picked ,Examined and assessed for this disease.

2.Random selection.

A sample in which each individual in the population has an equal chance of appearing is a random sample .Random ,here does not mean haphazard ,but it indicates the chance of the population unit being selected in the sample. So ,it is also called probability sampling.

Sampling Designs

There are different sample designs depending on the type and the nature of population and the objectives of the investigation. Some of the commonly used design are:

- 1.Simple random sampling
- 2.Systematic random sampling
- 3.Stratified random sampling
- 4.Cluster sampling
- 5.Multistage sampling

Simple Random Samples

Advantages

- Minimal of the knowledge of the population needed
- Easy to analyze data

Disadvantages

- Low frequency of use.
- Does not use researchers expertise.
- Larger risk of random error.

Stratified Random Sampling

- Population is divided into two or more groups called strata
- Subsamples are randomly selected from each strata

For example: if it is decide to know the prevalence of caries in different age groups, then the age groups from the strata and a random sample is to be chosen from each stratum.

Advantages

- ❖ Assures representation of all groups in sample population.
- ❖ Characteristics of each stratum can be estimated and comparison made

Disadvantages

- ❖ Requires accurate information on proportion of each stratum
- ❖ Stratified lists costly to prepare

CLUSTER SAMPLING

- The population is divided into subgroups (clusters) like families.
- A simple random sample is taken from each cluster



Cluster Samples

Advantages

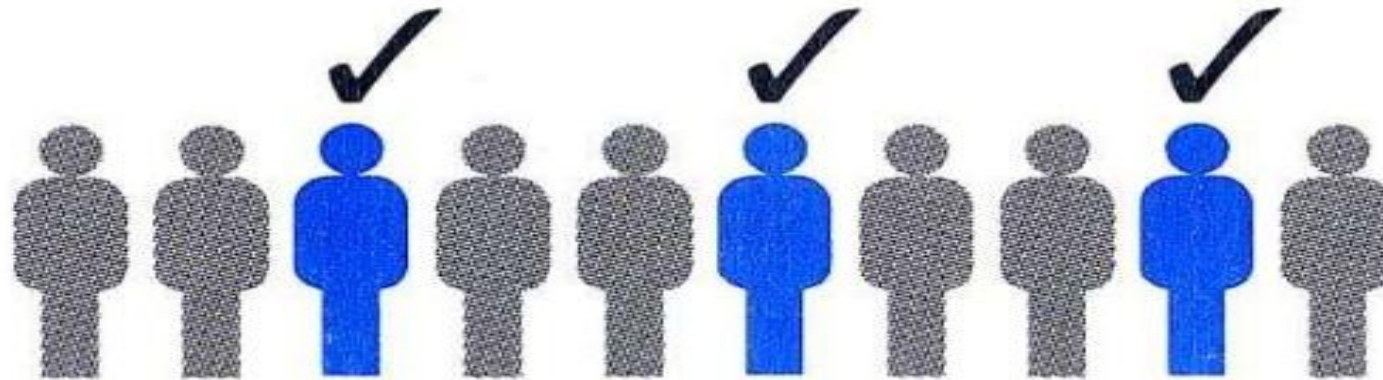
- ❖ Can estimate characteristics of both cluster and population

Disadvantages

- ❖ The cost to reach an element to sample is very high
- ❖ Each stage in cluster sampling introduces sampling error-the more Stages there are, the more error tend to be

SYSTEMATIC RANDOM SAMPLING

- ◉ Order all units in the sampling frame
- ◉ Then every n th number on the list is selected
- ◉ N = Sampling Interval



Systematic Random Sampling

Advantages

- ❖ Moderate cost, moderate usage.
- ❖ Simple to draw sample
- ❖ Easy to verify

Disadvantages

- ❖ Periodic ordering is required

MULTISTAGE SAMPLING

- Carried out in stages
- Using smaller and smaller sampling units at each stage

Primary
Clusters

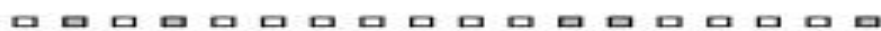


Secondary
Clusters



1

Simple Random Sampling within 5



Multistage Sampling

Advantages

- ❖ More accurate
- ❖ More effective

Disadvantages

- ❖ Costly
- ❖ Each stage in sampling introduce sampling error. The more stages there are the more error tend to be

Sample Size

The sample size need to be decided and calculated before selection and starting the study .Size of the sample is dependent on the statistical characteristics of the data to be collected.

- Larger the sample less the sampling variation i.e. Less chance of misleading results.

Most important point that the sample should be representative.

SAMPLING ERRORS




There are two types of errors that arise in sampling investigation:

1. The sampling errors which occur due to the sampling process, including:

- a. Faulty sampling design
- b. small size of the sample

2. Non sampling error arise due:

- a. Coverage error due to  non-response of the informant
non cooperation of the informant

b. Observational errors due to :

- ★ interviewers bias.
- ★ imperfect experimental technique.
- ★ interaction of both.

c. processing errors:

 Due to errors in statistical analysis



Thank you for your kind attention