

Pharmaceutical Statistics

Lecture 1

Introduction to statistics

Basic concepts of statistics

Prepared and presented by

Dr. Muna Oqal

Basic concepts of statistics

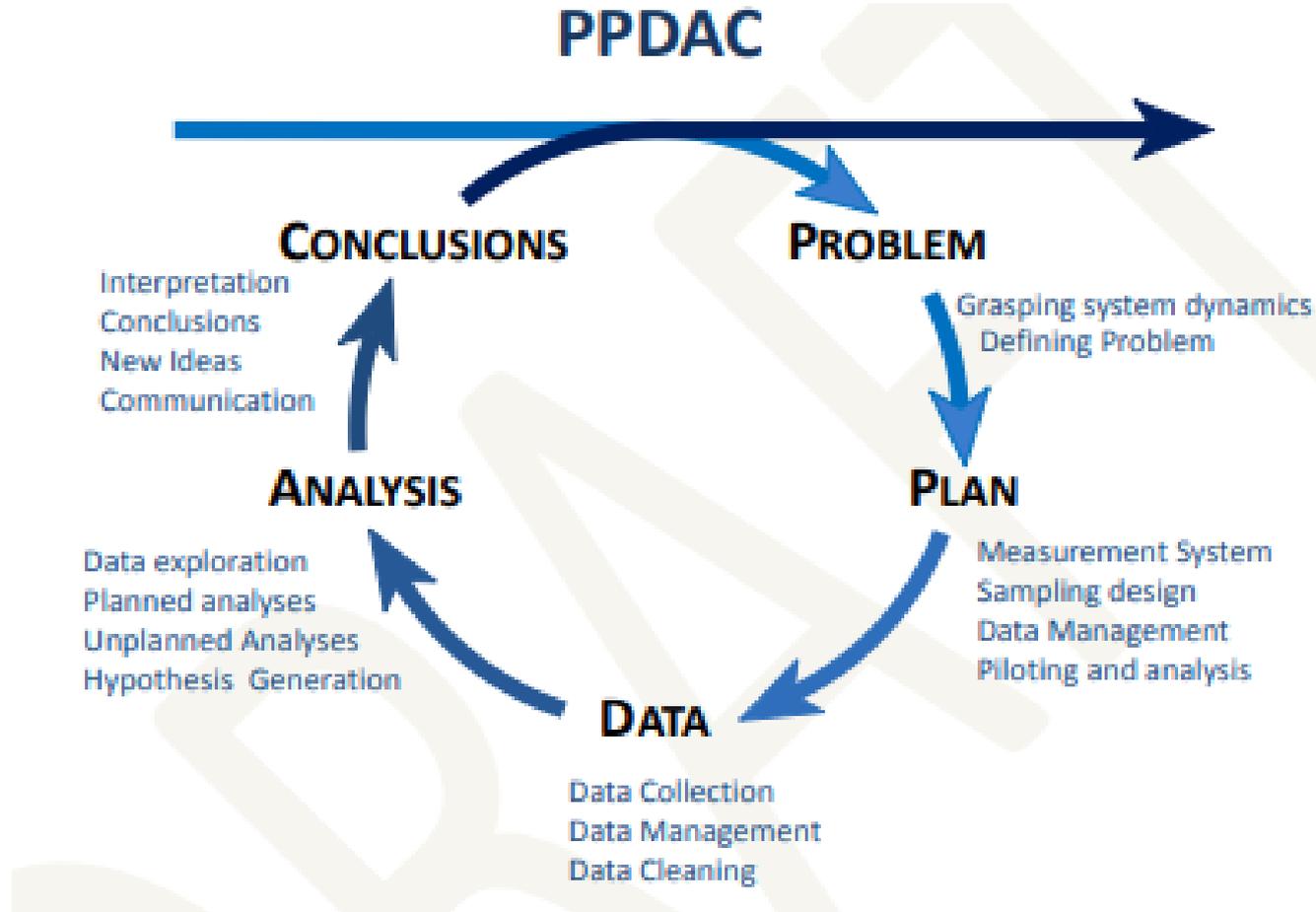
What is statistics?

- Statistics is a field of study concerned with:
 1. Collection, organisation, summarisation and analysis of data.
 2. Drawing of inferences (conclusion) about a body of data when only part of data is observed (inferential statistics).
- Statistics provides a way of offering critical guidance in producing trustworthy analyses and predictions.

Biostatistics

- The tools of statistics are employed in many fields: Business, Pharmacy, Medicine, Education, Psychology,....etc.
- When the analysed data are derived from the biological science, Pharmacy, and Medicine, we use the term Biostatistics to distinguish this particular application of statistical tools and concepts.

The Statistical Inquiry Cycle



Pharmaceutical Statistics

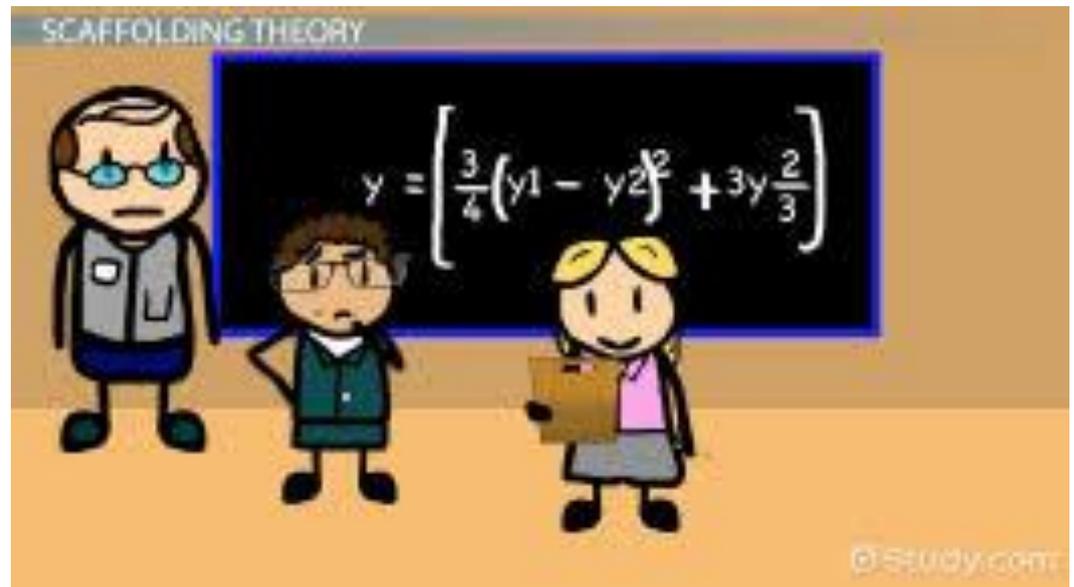
- The pharmaceutical statistics is an application of statistics to matters concerning the pharmaceutical industry. This can be from issue of design the experiments, to analysis of drug trials, to issues of commercialisation of a medicine.

Branches of Statistics

- **Descriptive Statistics:** is a branch of statistics that involves the collections, organisation, summarisations, and display of data.
- **Statistical Inferences (Inferential Statistics):** is a branch of statistics that involves using a sample to draw conclusions (inferences) about a population.

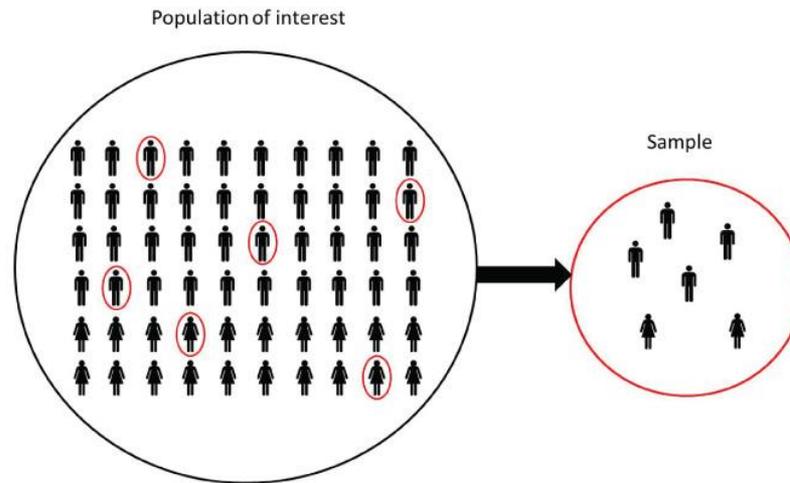
Vocabulary of Statistics

- Population:
- Sample
- Data
- Variables
- Parameter
- Dependency



Vocabulary of Statistics

- **Population**: A population is an entire group, collection of all items or counts that are of interest of a particular study which we want to characterise.
- **Sample**: A sample is a collection of observations on which we measure one or more characteristics.



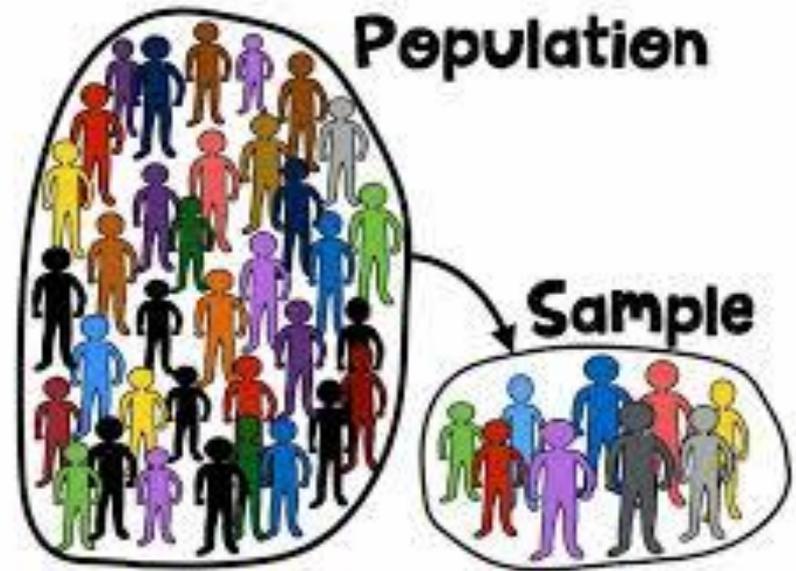
- The sample should be selected to represent the whole population.
- The sample should be **representative** of the population and **unbiased**.
- When the sample does not represent the population, it is called **unrepresentative**.

Vocabulary of Statistics

- Descriptive measures that describe a **population** are called a parameter.
- Descriptive measures that describe a **sample** are called a statistics.
- Parameters: Quantities (numerical measurement) describing some characteristic of a population.
- This means the parameter tells us something about the whole population.
- In statistics, we draw a conclusion about a population on the basis of information contained in the sample that have been drawn from the population

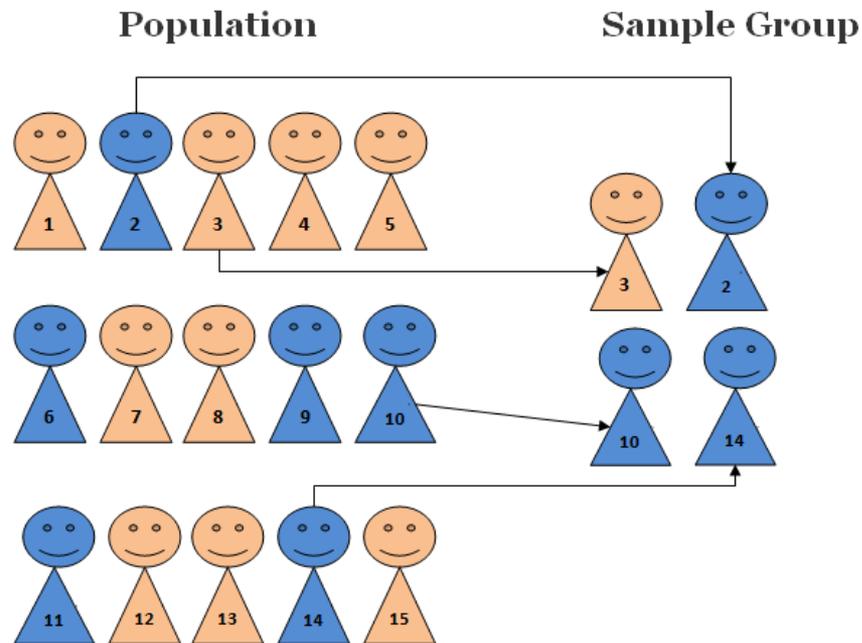
Why we study a sample rather than a population?

- Less cost
- Preservation from loss
- Save time
- More accuracy
- Population is so huge



Choosing samples

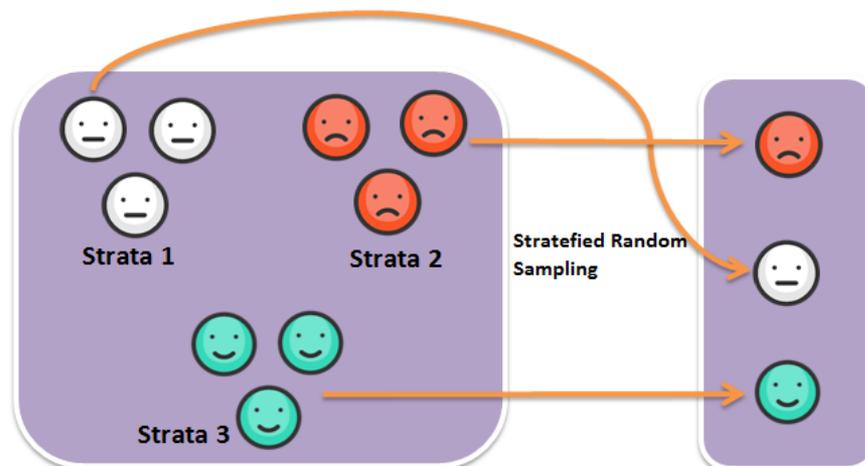
- **Simple Random Sample (SRS):**
 - Simplest type of scientific samples.
 - Every individual in a sample is chosen randomly and entirely by chance.



Choosing samples

- **Stratified sampling**

- The population is divided into separate groups (strata) with similar characteristics and then individuals or objects can be randomly selected from each group according to its % in the whole population.
- For example, in a study we may wish to ensure a certain percentage of smokers (25%) are represented in both the control and experimental groups in a clinical trial (n=100 per group).

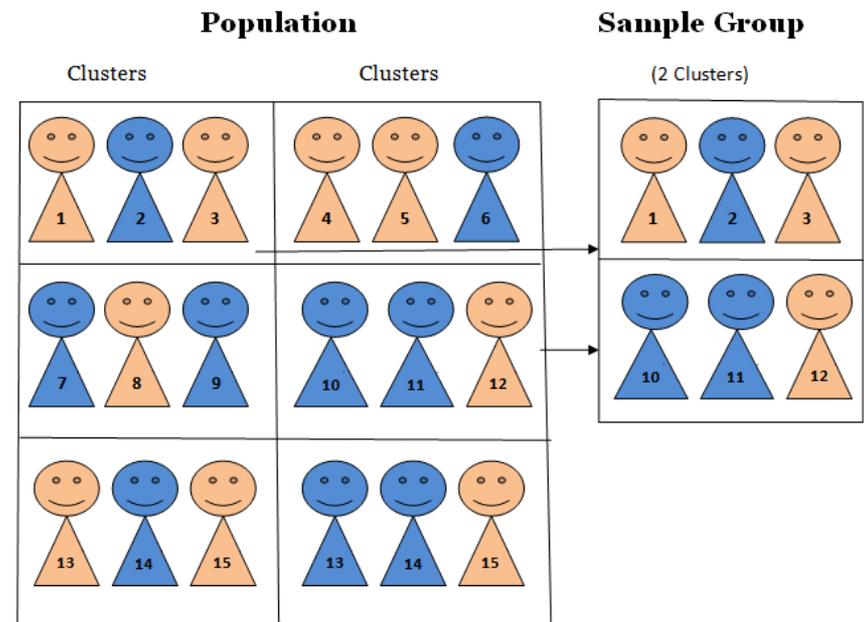


Choosing samples

- **Cluster sampling "multistage" sampling**

- Is employed when there are many individual "primary" units that are clustered together in "secondary", larger units that can be subsampled.

For example, assume that 150 containers of a bulk powder chemical arrive at a pharmaceutical manufacturer and the quality control laboratory needs to sample these for the accuracy of the chemical or lack of contaminants. Rather than sampling each container we randomly select ten containers. Then within each container of the ten containers we further extract random samples (from the top, middle bottom) to be assayed.



Choosing samples

- **Systematic sampling (SYS):** a sample is selected by listing a population sequentially and choosing members at regular intervals or according to a set schedule or plan.



In this case, every third person is systematically selected

Vocabulary of Statistics

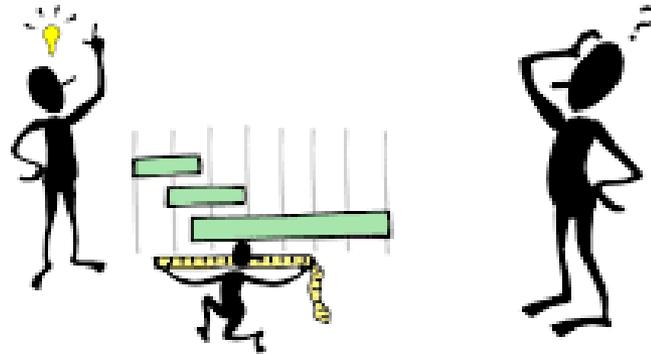
- **Data**: the raw material of statistics. They are derived from observations or measurements. We may define data as a numbers.
 - There are two kinds of numbers that we used in statistics:
 - Numbers generated from measurements (e.g. temperature)
 - Numbers generated from a process of counting (e.g. # of patient discharged from hospital).
 - Data may be generated or obtained by one or more of the following sources:
 - ✓ Surveys
 - ✓ Experiments
 - ✓ Routinely kept records
 - ✓ External resources

Vocabulary of Statistics

- **Variables**: a characteristic monitored, capable of taking different values (diastolic BP, height of students,...etc).
- Type of variables according to their measurability:
 1. Qualitative variables
 2. Quantitative variables



What is a Variable?



Variables

- **Qualitative variables**

- Variables that can be placed into distinct categories (non-numerical), according to some characteristic or attributes.
- Not measurable, categorical characteristics
- Examples: Sex (male or female), Blood groups (A, B, O, AB)...etc.

- **Quantitative variables**

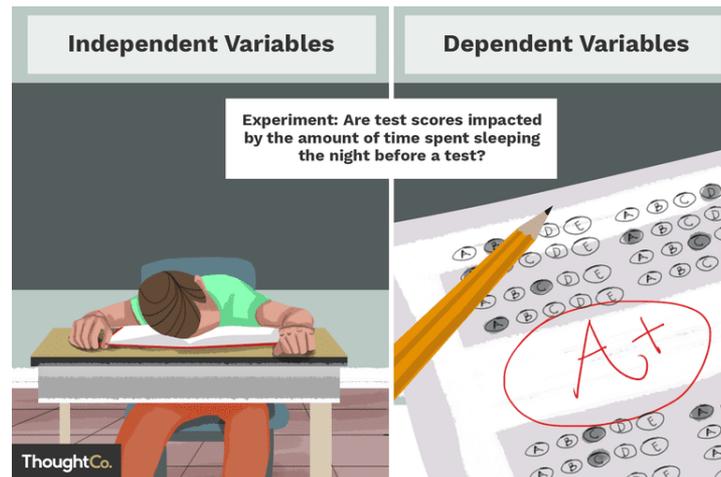
- Variables that are numerical in nature and can be ordered or ranked, which generated usually from counting or measuring something.
- Measurable.
- For examples the variable age is numerical and the people can be ranked in order according to their age.

Statistical Studies and Variables

- There are different approaches to classify statistical studies, examples of these studies are:
 - **Observational study:** In which the researcher is merely observes the effect pf a risk factor, diagnostic test or what is happening or what has happened in the past and tries to draw a conclusion based on these observations (e.g. cohort studies, and case control studies). No intervention to the sample
 - **Experimental study:** In which the researcher manipulates one of variables (or introduce an intervention) and tries to determine how the manipulation (or intervention) influences other variables. Example is Randomized Control Trials (RCT).

Vocabulary of Statistics

- **Dependency**: An independent variable, sometimes called an experimental or predictor variable, is a variable that is being manipulated (controllable and can be specified in advance before conducting the experiment) in an experiment in order to observe the effect on a dependent variable, sometimes called an outcome variable.
- For examples: for the effect of the concentration of amylase inhibitor on enzyme activity, inhibitor concentration is the independent variable and enzyme activity is the dependent variable.



Computer and Calculators

- In the past, statistical calculations were done with pencil and paper, but now computers and calculators do all numerical calculations.
- The most commonly used statistical package are:
 - GraphPad prism
 - SPSS
 - STATA
 - SAS
 - MINITAB
- As students, you should know that even the computer and calculators can do perfectly all the numerical calculations in effectively time-manner. But you are still responsible for understanding and interpreting each statistical concept.