

Principles of Statistics

The science of Statistics **deals with the collection, analysis, interpretation, and presentation of data in such a way** that meaningful conclusions can be drawn from them. Inferential statistics is a scientific discipline that uses mathematical tools to make forecasts and projections by analyzing the given data.

In generally, statistics is the methodology for collecting, analysing, interpreting and drawing conclusions from information and collected data.

There are two types of Statistics:

. • **Descriptive Statistics (Exploratory):** involves organizing and summarizing data. is concerned with summary calculations, **graphs, charts and tables** about a given data.

• **Inferential Statistics(Confirmatory):** is a method used to generalize from a sample to a population. sometimes called analytical statistics

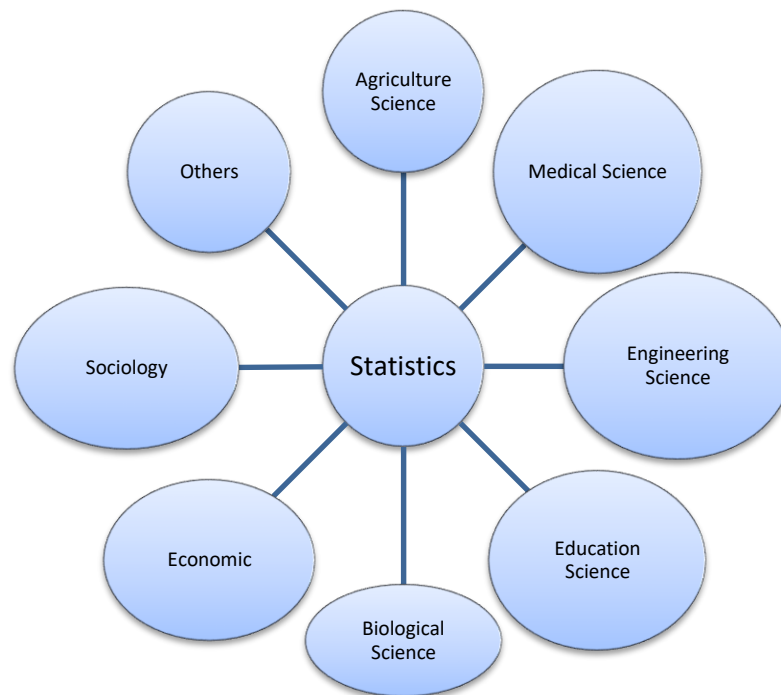


Figure: Relationship between statistics and other sciences

Functions of Statistics

1. Definiteness
2. Condensation
3. Comparison
4. Formulating and testing of hypothesis
5. Prediction or Forecasting
6. Planning and Decision

Key Terms (Common Expressions)

Variable: A variable, denoted by capital letters such as X and Y, is a characteristic of interest for each person or thing in a population. Variables may be numerical or categorical. **like student's length, student's weight, Tree's height, number of flowers per plant...etc.**

Types of variables:

1-Qualitative (or categorical) variable: It is a variable which can't be measured in quantitative form, but can only be identified by name or categories

E.g. types of color, types of birth (single, twin), sex (males and females).

1-eye's colour (Blue, Black, etc.)

2. Color of a person's hair (black, grey, red, brown, etc.)

3. Gender of child (male, female)

4. Cause of death of new born (congenital malformation, asphyxia, . . .)

2- Quantitative variable: it is a variable in which its observations are in numerical measurements like, blood pressure value, Serum cholesterol, plant height, germination %. Number of oranges /tree, number of leaves /plant, number of branches /tree. Quantitative variable or characters may also be classified to:

a. Continuous Variables.

Continuous variable: Means that all possible values are observing on continuous scale like the blood pressure of 5 students are (11, 12,13,14, and 15 cm Hg).

b. Discontinuous (discrete) Variables. Means that all possible values are not observing on a continuous scale. like the blood cholesterol of 5 students are (179, 181, 189, 191, 190). Or the organic matter % content of 5 forest soils are :3 ,4 ,6 ,7 and 8.

Population and Sample:

Population: Is any specific collection of objects of interest (The collection of all possible outcomes)

There are two types of population:

1-Finite population: It is a population which its number can be count.

2-Infinite population: It is a population which its number cannot be count.

Sample: It is any subset or sub collection of the population, including the case that the sample consists of the whole population (a sub-set of a population). A sample that represent the characteristics of population called representative sample.

Data: Numbers or measurements collected. One observation called datum or measurement.

Parameter: is a characteristic of a population. Are the measures determined from populations such as **population mean (μ)**.

Examples: The average height of all Britons.

Sources of Data: primary and secondary

1-Primary Data: a data collected by the user directly from the source.

Methods of collection

Personal Interview face-to-face, Telephone Focus Group discussion (FGD)

2-Secondary Data: a data gathered or compiled from published and unpublished sources. **From** journals, reports, government publications, publications of professionals and research organizations.

E.g.- CSA: Central statistics agency

DHS: the demographic and Health Survey

HDS: Health and Demographic Surveillance

Sample mean, is pronounced (\bar{x})

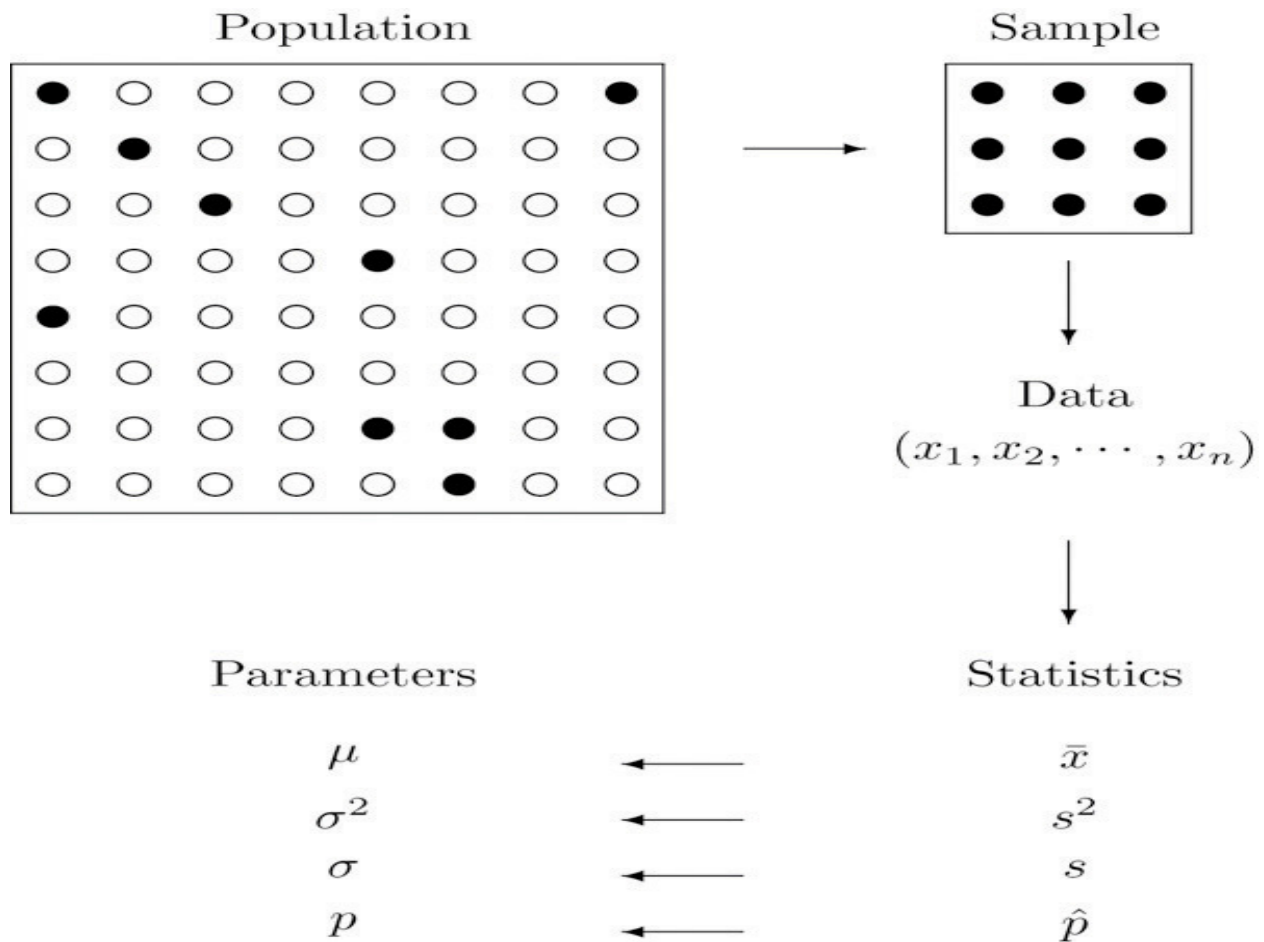
n = Sample size N = Population size

μ = Population mean ,is pronounced (μ)

M_o = Mode M_e = Median

S^2 = Sample variance σ^2 = Population variance

S = Sample standard deviation



$\sigma_{\bar{x}}$ = Population standard error $S_{\bar{x}} = SE$ = Standard error or standard deviation of mean

C.V. = Coefficient of variation.

P = Probability

n! = Factorial of n.

SS = Sum of squares.

Σ = Summation (sigma)