

Overview of Elementary Statistics

Statistics is an area of Mathematics that is concerned with the collection, organization and description of data (Descriptive) and the interpretation, analysis and prediction of data (Inferential).

Statistics involves activities of two types of investigations observational studies and true experiments.

Observational studies are naturalistic in which no variables are manipulated by the investigator.

True experiments are investigations to find effects of changing one variable on another variable.

Statistics is normally divided into two categories:

Descriptive Statistics is involved with collection, organization and description of (data) information.

Inferential Statistics is involved with interpretation, analysis and prediction of (data) information.

Statistics is an important discipline since it increases or contributes to man's collection of knowledge.

There are many ways man uses to establish knowledge and a specific method is the Scientific Method.

1. Define a clear and precise statement of the problem or phenomena to be studied.
2. Collect or gather all available information about the subject that is to be investigated.
3. Form a hypothesis or prediction about the subject which is a preliminary explanation.
4. Design an experiment to test the truth of the preliminary hypothesis or prediction.
5. After experiments a conclusion is arrived as to the truth of the hypothesis or prediction.

It does not matter whether hypothesis is proved to be true or false since collected knowledge is foremost.

One vital aspect of Scientific Method is that it can be used to validate any experiment by another investigator.

Elementary Statistical Terms

Population:	A population is a set of all possible subjects, objects or scores that an experimenter is concerned with during an experiment or study.
Sample:	A samples is part of or a subset of a population that an experimenter uses instead of a large experiment during experiment or study.
Data:	Data is actual observations or measurements that are collected by the investigator during a study or experiment from a population or sample
Variable:	A variable is any characteristic, object, or quantity that may take on different values as a result of actual manipulation or dependence.
Independent Variable:	The independent variable in an experiment or study is the variable the investigator systemically varies during an experiment or study.
Dependent Variable:	The dependent variable in an experiment or study is the variable the investigator systemically measure during an experiment or study.
Parameter:	A calculated value which has been arrived at by using raw data of a population.
Sample:	A calculated value which has been arrived at by using raw data of a sample.

Measurement (Data Types) Scales

- Nominal:** A nominal scale or data is the lowest form of measurement used in statistical investigations. It is likely to be qualitative in nature as to a quantitative value or measurement.
- Examples: Choices, Preference, Attributes...
- Ordinal:** An ordinal scale or data is the next higher measurement used in statistical investigations. It has the property of relative magnitude which allows for: $A < B$, $A = B$, $A > B$
- Examples: Positions and Rankings...
- Interval:** An interval scale or data is the next higher measurement used in statistical investigations. It possesses properties of magnitude and spacing between adjacent data but has no Absolute Zero.
- Examples: Rulers and Temperatures...
- Ratio:** A ratio scale or data is the highest type of measurement used in statistical investigations. It has all attributes of lower scales and Absolute Zero which allows for proportional comparisons.
- Examples: Any scale with above properties.

Measure of Central Tendency

- Definition:** Numbers or indicators that express how scores seem to bunch together.
- Mode:** The most frequent score or the score that happens the most times.
- Median:** The physical middle score or the score in the middle of ranking.
- Mean:** The arithmetic average or $(\text{Sum of all scores} / \text{Total number of scores})$

Measure of Central Tendency

- Definition:** Numbers or indicators that express how scores spread out or scatter.
- Range:** The difference between the High & Low scores: $(\text{High Number} - \text{Low Number})$
- Variance:** The quotient of the deviations (differences) squared & total number of scores.
- Standard Deviation:** The square root of the variance. (Simply stated: Average Deviation)

Classical _ Statistical _ Formulas

$$\bar{X} = \sum_{i=1}^n X_i / N \quad \mu_X = \sum_{i=1}^n X_i / N$$

$$s = \sqrt{\sum_{i=1}^n (X_i - \bar{X})^2 (f) / N - 1} \quad \sigma_X = \sqrt{\sum_{i=1}^n (X_i - \bar{X})^2 (f) / N}$$

$$SS = \sum_{i=1}^n X_i^2 - (\sum_{i=1}^n X_i)^2 / N$$

$$s = \sqrt{SS / N - 1} \quad \sigma_X = \sqrt{SS / N}$$

$$\text{Median} = LB + ((Nf / Cf)(CS))$$

$$\text{Median} = UB - ((Nf / Cf)(CS))$$

$$Z = (X - \bar{X}) / S = \frac{x - \mu}{\sigma}$$