

Chapter 1

Introduction to probability and Statistics

1.1 Definition of STATISTICS

- **Statistics** is a science deals with collect, organize, analyze, and draw meaningful inferences from data, which lead to good decisions.
- **Statistics** is a science that helps us make better decisions in business and economics as well as in other fields.

1.2 Classification of Statistics

- **Descriptive Statistics** helps to describe a given set of data without going beyond that data
 - It consists of collection, organization, summarization and analysis of the data
- **Inferential Statistics** helps to make inference/conclusion about a population based on the selected sample
 - it consists of predict and forecast values of population parameters; test hypotheses about values of population parameters; make decisions

1.3 Stages in Statistical Investigation

- **Stage 1: Data collection**
- **Stage 2: Organization and presentation of data**
- **Stage 3: Data analysis**
- **Stage 4: Interpretation of the result**

■ **Data collection**

The process of measuring, gathering, assembling the raw data up on which the statistical investigation is to be based.

Data can be collected by using survey. Survey can also be done in different methods, three of the most common methods are:

- Telephone survey
- Mailed questionnaire
- Personal interview.

■ **Organization and presentation of data**

The process of organization, classification, compilation, and summarization of data to present it in a meaningful form.

Example;- graphs, charts and tables.

Analysis of data

The process of extracting relevant information from the summarized data, mainly through the use of elementary mathematical operation.

Inference of data

By using analysis of the data conclusions are formed and inferences made.

1.4 Definition of some basic terms

- A **population** consists of the set of all measurements/elements under study.
- A **sample** is a subset of the measurements selected from the population.
- A **census** is a complete enumeration of every item in a population.
- **Sampling** is the process of taking a sample from a population

- **Parameter** is a statistical measure obtained from a population data
- **Statistic** is a statistical measure obtained from a population data
- **Variable** is a characteristics under study that assumes different values for different elements

1.7 Types of variables

- **Quantitative variables:** are quantifiable

- **Discrete:** assumes countable/countably infinite values.

Example: Number of students in a class; number of cars in a parking lot etc

- **Continuous:** assumes values in intervals

Example: weight, height, length, temperature etc

- **Qualitative / Categorical variables:** are non quantifiable

Example: color, nationality, sex, ...

1.5 Application of Statistics

- **It is applied in any field of study which seeks quantitative evidence**
- **For instance,**
 - **To compare breaking strength of two products**
 - **For quality control of products in a given production process**
 - **To compare the improvement in yield due to application of fertilizer, pesticide....**

1.6 Limitation of Statistics

- **It does not study qualitative characteristics directly**
- **It doesn't deals with a single individuals but deals with aggregate of facts**
- **Statistical findings are approximate**
- **It is sensitive to misuse**

Scales of measurement

Property of measurement scale

Order

- The property of order exists when an object that has more of the attribute than another object, is given a bigger number by the rule system. This relationship must hold for all objects in the "real world".

The property of ORDER exists

When for all i, j if $O_i > O_j$, then $M(O_i) > M(O_j)$.

Distance

The property of distance is concerned with the relationship of differences between objects.

The property of DISTANCE exists, for all i, j, k, l

If $O_i - O_j \geq O_k - O_l$ then $M(O_i) - M(O_j) \geq M(O_k) - M(O_l)$.

Fixed Zero

- A measurement system possesses a rational zero (fixed zero) if an object that has none of the attribute in question is assigned the number zero by the system of rules.
- The property of FIXED ZERO exists if $M(O_0) = 0$.
- The property of fixed zero is necessary for ratios between numbers to be meaningful.

1.8 Types of Scales of Measurement

- **Nominal Scale** - Nominal scales are measurement systems that possess none of the three properties stated above.
 - **used for grouping or classification**
 - ✓ Color, sex, religion, ...
 - ✓ +, -, *, / are impossible
 - ✓ Comparison is impossible
- **Ordinal Scale** - Ordinal Scales are measurement systems that possess the property of order, but not the property of distance. The property of fixed zero is not important if the property of distance is not satisfied.

- **used for grouping & ordering**

- ✓ **Economic status (poor, medium, higher), military ranks...**
- ✓ **+, -, *, / are impossible**
- ✓ **The magnitude b/n the values is not clearly known**

- **Interval Scales**

Interval scales are measurement systems that possess the properties of Order and distance, but not the property of fixed zero.

- ✓ **The magnitude b/n the values is clearly known**
- ✓ **has arbitrary zero value**

- ✓ **Temperatures ($^{\circ}\text{F}$, $^{\circ}\text{C}$)**
- ✓ **+ & - are possible but * & / are impossible**

• **Ratio Scale** - Ratio scales are measurement systems that possess all three properties: order, distance, and fixed zero.

- ✓ **Has true zero point**
- ✓ **+, -, *, / are possible**
- ✓ **salary, weight, height,....**