



كلية العلوم

قسم الذكاء الاصطناعي

Department of Artificial Intelligence

Subject : Basic Concepts

Class: 1st

Lecturer: Asst.Lect Ali Hussein Jaber

Lecture: (1)

Definition of Statistics

Statistics is the scientific discipline that focuses on the collection, organization, summarization, presentation, and analysis of data to derive meaningful conclusions and make informed decisions.

Statistics is generally divided into two main branches:

1. **Descriptive Statistics**

This branch deals with methods for describing a specific set of data. It involves techniques for collecting data in numerical form, organizing and summarizing it, presenting it effectively, and calculating various statistical measures to describe the dataset.

2. **Inferential Statistics**

This branch focuses on drawing conclusions or making inferences about the source population from which the data were collected. Inferential statistics includes two primary areas:

2.1 Estimation: This involves finding estimated values that approximate the true values of the population. These estimates can be either **point estimates**, providing a single value, or **interval estimates**, giving a range of values.

2.2 Hypothesis Testing: This involves formulating initial assumptions (hypotheses) about the phenomenon under study and making decisions about whether to accept or reject these hypotheses based on the collected data.

Variables and Their Types

A **variable** is any characteristic or phenomenon that exhibits variation among its elements. Variables are commonly denoted by symbols such as Y , X , Z , etc. Variables can be classified into:

1. **Qualitative (Categorical) Variables:**

These are characteristics that cannot be measured numerically. Examples include eye color (blue, black, brown), social status (rich, middle class, poor), and gender (male, female).

2. **Quantitative (Numerical) Variables:**

These are characteristics that can be measured using numbers. Examples include height, weight, age, and the quantity of a product. Quantitative variables are further divided into:

2.1 Continuous Variables: These can take any numerical value within a given range. For instance, if students' heights at a university range between 130.5 cm and 170 cm, we can write: $130.5 \leq Y \leq 170$

2.2 Discrete Variables: These take distinct, separate values. For example, the number of family members in four households might be 2, 3, 4, and 5, or rolling a die can result in 1, 2, 3, 4, 5, or 6.

Population and Sample

1. **Population:** This refers to the complete set of values or observations that a variable can take. For example, if our study focuses on the heights of students at a particular university, the population consists of the heights of all students at that university.

1.1 Finite Population: A population whose elements can be fully counted, e.g., the heights of students at Al-Mustaqbal University.

1.2 Infinite Population: A population whose elements are difficult or impossible to count, e.g., the population of a certain fish species in the Tigris River or the number of bacteria in a field.

2. **Sample:** A sample is a subset of the population, selected using a specific method.

Studying the entire population may be difficult, time-consuming, or costly. Therefore, researchers often study a sample and use its characteristics to infer properties of the original population.