جامعة المستقبل
AL MUSTAQBAL UNIVERSITY

جامعة المستقبل
كلية العلوم – قسم الانظمة الطبية الذكية

Lecture: (1)

Subject: Statistics and probability
Class: Second
Lecturer: Prof. Dr. Mehdi Ebady Manaa

CHAPTER ONE
Introduction to Probability Theory and Statistics

Statistics in mathematics is a branch of applied mathematics that involves collecting, describing, and analyzing data and then drawing conclusions from the available data. Therefore, statistics can be defined as the science that is concerned with collecting data, converting it into a numerical form, and then organizing, arranging, and analyzing it to arrive at specific results that help explain a particular phenomenon.

Basic concepts in statistics.

- **Data:** which is one of the most important basic concepts of statistics. It represents the signs or observations observed in the study population, and is expressed in numbers. Data can be divided into **two parts:**
  1. **ungrouped data :** It is the primary data that was collected and not classified
  2. **Grouped data :** It is data organized in a frequency distribution table.

- **A Statistical variable:** is any characteristic that can be measured or monitored. These variables vary among themselves depending on the field in which they are used.

  Types of statistical variables, according to the nature of the data:

1. **Descriptive variables:**
   These are variables that can be described verbally. It is divided into **nominal** variables and **ordinal** variables.

2. **Quantitative variables:**
   These are variables that can be measured and recorded digitally, such as weight, height, and age, and can be divided into:

- **Continuous variables:** Variables that have any numerical value within a specified range. If we assume that the heights of a group of people range between 130 cm and 170 cm, we can write the variable in the form:

$$130 \leq X \leq 170$$

- **Discrete variables:** They are variables between which there is a clear distinction that distinguishes them, such as gender, either male or female. The number of things is always discrete variables, such as the number of children in the family, the number of students in the class, and the number of teachers in the school.

➢ **Population**: It is all the values or items that a variable can take.

The statistical population is divided into:

1- Limited: which includes a limited number of individuals, such as the number of 100 students in the first semester of the year 1432 AH.

2- Unlimited: It is the one in which the number of individuals is infinite (unlimited), such as the number of 100 students for the next ten years (assuming the continuation of the course).

In most cases, it is difficult or impossible to observe the data of all members of society, such as research conducted to determine the rate of illiteracy محو الامية in a country or city, and research that aims to count the harvested wheat grains. To overcome this, a part of society can be selected, called a sample.

➢ The **Statistical sample** is known as:

It is a subset of the statistical population, chosen so that it faithfully represents the statistical population from which it was drawn.

Reasons for the necessity of studying a sample instead of a population:

1. The **difficulty** or impossibility of examining the entire community due to:

   - ✓ Its large size, as in estimating the fish wealth in a community.
   - ✓ The inspection may be destructive to the units, as in examining the lifespan of bulbs produced by a specific factory.
   - ✓ The examination may be harmful to the units, such as examining the patient's blood.

2. **Costs** and capabilities (examining the entire community costs a lot of effort and money) in showing the results.

3. **Accuracy** of data and information due to the possibility of using highly qualified and trained people.

Example

Simple Statistical Analysis fo sample of data

Objective: to Perform a statistical study to analyze some sample data

```
import numpy as np

# Step 1: Sample patient weight data (in kilograms)
weights = [65, 72, 78, 85, 90, 55, 60, 75, 80, 70]

# Step 2: Calculate basic statistics
mean_weight = np.mean(weights)  # Average weight
median_weight = np.median(weights)  # Middle value
std_dev_weight = np.std(weights)  # Standard deviation
min_weight = np.min(weights)  # Minimum weight
max_weight = np.max(weights)  # Maximum weight

# Step 3: Print the results
print(f"Mean Weight: {mean_weight:.2f} kg")
print(f"Median Weight: {median_weight:.2f} kg")
print(f"Standard Deviation: {std_dev_weight:.2f} kg")
print(f"Minimum Weight: {min_weight} kg")
print(f"Maximum Weight: {max_weight} kg")
```