الفصل الأول

الإحصاء ومفاهيمه الأساسية

1-1Introduction

Statistics is a branch of applied mathematics that involves the collection, description, analysis, and inference of conclusions from quantitative data. **Statisticians**, people who do statistics, are particularly concerned with determining how to draw reliable conclusions about large groups and general events from the behavior and other observable characteristics of small samples.

1-2Historical Background of Statistics

Statistics is an ancient science that has been linked since the beginnings of civilizations to organizing human life, counting resources, and making decisions. The development of statistics over time can be summarized as follows:

1. Ancient Times

- **Ancient Egypt:** Egyptians used statistics to count the population, record agricultural lands, and calculate taxes
- **Ancient China:** The Chinese relied on collecting population and agricultural data to distribute resources and manage the state.
- **Ancient Greece:** Greeks used statistics to count soldiers, organize armies, and study populations, births, and deaths.

2. Middle Ages

- In Europe, statistics were used to estimate military strength and collect taxes. Cities also recorded population and disease data to monitor public health.
- In the Islamic world, information about populations, agriculture, and trade was documented. Statistics are also mentioned in the Quran in contexts related to organizing community resources.

3Contemporary Era

With the digital revolution, statistics has become essential for big data analysis, artificial intelligence, weather forecasting, business management, medicine, and engineering.

Today, statistics is a vital tool for understanding the world and making data-driven decisions in all areas of life.

1-3 Statistics

Statistics is the study and manipulation of data, including ways to gather, review,0analyze, and draw conclusions from data. It is scientific tool that can be used in several areas such as engineering, management, agriculture and computer science and networking.

1-3-1There are two types of statistics are known:

- **Descriptive statistics:** which describes the properties of sample and population data · Collecting, Organizing, Summarizing, Presenting Data.

- **Inferential statistics :** which uses the properties to test hypotheses and draw conclusions· Making Inferences, Hypothesis Testing, Determining Relationships, Making Predictions.

1-4Statistics Concepts

Statistical Term (Observation): It is the numerical data related to a specific subject.
Therefore, the observation is the primary material used in the field of statistics.

Statistical Population: The statistical population refers to the set of statistical units intended to be studied and accurately understood, which share the main characteristic of interest to the researcher.

Classification of Statistical Populations by Type:

- **Finite Populations:** These are populations whose total units can be known, such as the number of first-year students in the Faculty of Computer Science and Informatics.
- **Infinite Populations:** These are populations whose units cannot all be known, such as the number of stars in the sky.

Sample: A sample is a subset selected from the population using a specific statistical method so that it accurately represents the population from which it was drawn, and it is relied upon in the study instead of the entire population.

Random Variables: Variables whose values are determined by chance.

- **Discrete Variables:** Variables that take values that can be counted.
- **Continuous Variables:** Variables that can take any value between a given minimum and maximum.

Types of Variables:

- **Qualitative Variables:** Deal with descriptions; data can be observed but not measured (e.g., colors, appearance, and others).
- **Quantitative Variables:** Deal with numbers; data can be measured (e.g., length, height, area, volume, weight, speed, time, temperature, cost)

Parameter: A parameter is defined as a statistical constant that is often unknown and is estimated using a specific method.

1-5 Stages of Statistical Research

- Defining the Research Problem
- Data and Information Collection
- Data Classification, Tabulation, and Presentation
- Data Analysis and Predictions
- Interpretation and Decision Making

1-6 Data Collection Sources

Data collection sources are divided into two main types:

1. Primary Sources

These are data collected for the first time directly from the field or from their original source. They are characterized by being recent and suitable for the purpose of the research.
The most important methods include:

- Observation
- Interview
- Questionnaire
- Experiments

- Records and direct events

2. Secondary Sources

These are data previously collected by other researchers or institutions and then reused in the current research.
The most common forms include:

- Books and scientific references
- Official reports
- Government statistics
- Articles and journals
- Ready-made databases

1-7 Data Collection Method

- **1. Complete Enumeration Method:**
- This method involves counting all the statistical units that make up the population under study.

- **2. Sampling Method:**
- This is the process of selecting a part of the statistical population to infer the characteristics of the entire population by generalizing the results obtained from the sample.

1-8 Types of Errors

1. Bias Errors (also called Systematic Errors):

These errors result from various sources, such as errors in the design of the research or experiment, technical errors during data collection, or computational errors that occur while processing the collected data.

2. Sampling Errors (also called Random Errors):

These errors arise from random differences between the units of the population included in the sample and those units not selected, which leads to variations caused by chance during sample formation.

Note: To minimize random errors:

1. The sample must be selected using an appropriate method.
2. The sample size should be suitable for the size of the population.

1-9Statistical Symbols

< less than

> more than

≤ less than or equal

≥ more than or equal

|| absolute value

Σ summation

fi : frequency

n!: factorial

$\bar{y}$: mean

Me: median

Mo: mode

Ei : event

P : probability

∞: infinity