

Chapter 01

General Overview

Biostatistics For the Health Sciences

Dr. Moustafa Omar Ahmed Abu-Shawiesh

Professor of Statistics



Introduction

Statistics is a science whereby inferences (conclusions) are made about specific random phenomena on the basis of relatively limited sample material.



Important Concepts

➤ **Mathematical Statistics**

Concerns with the development of new methods of **statistical inference** and requires detailed knowledge of abstract mathematics for its implementation.

➤ **Applied Statistics**

Involves application of **mathematical statistical methods** to specific subject areas such as economics, psychology, public health and Medicine.

What is Statistics?

Answer

Statistics is a field of study concerned with the **collection, organization or presentation, summarization** and **analysis** of **data** for the purpose of drawing conclusions (**inferences**) and making decisions about **a population** based on only a part of the data (**Sample**) taken from that population.

Definition: Biostatistics (Biometry)

It is a branch of **applied statistics** that applies statistical methods to medical and biological problems.

Question: Why do we need to know and study **Biostatistics**?

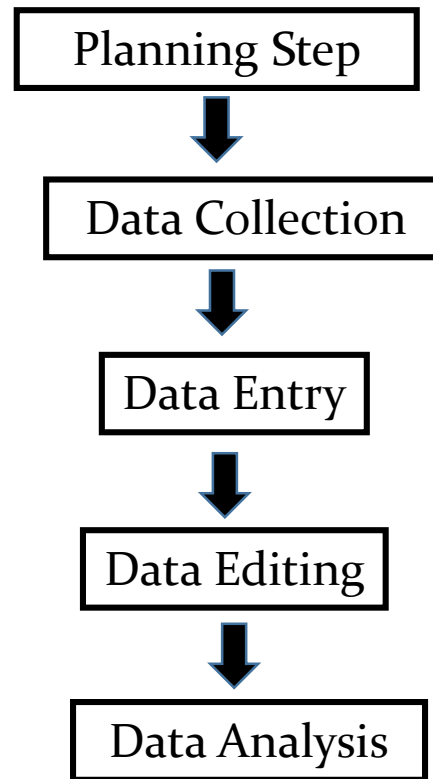
Answer

- Standard statistical methods may not necessarily be applicable for all studies.
- New bio statistical methods are developed by biostatisticians.

The Role of Biostatistics in Medical Research

A good way to learn about **biostatistics** and its role in the research is to follow the study from its inception at the planning stage to its completion, which usually occurs when the results of the study are published.

Statistical Study Steps



After completing the data-collection, data-entry, and data-editing phases, we were ready to look at the **results of the study**.

Planning Step: At this stage we had to make several important decisions, each of which proved vital to the success of the study.

Data Analysis: At this stage, data obtained from the study can be summarized using **descriptive statistics**.

Descriptive Statistics

Consists of methods for organizing, displaying, and describing data by using tables, graphs, and summary measures.

Descriptive material can be **numeric** or **graphic**:

- If numeric, data can be tabulated or presented as a frequency distribution.
- If graphic, data can be summarized pictorially.

Note that: The Choice of **numeric** or **graphic descriptive statistics** is dependent on type of distribution of data.

1. **Continuous data:**

- where there are infinite number of possible values (e.g., weight, age, blood pressure measurements).
- means and standard deviations may be used to describe data.

2. Discrete data:

- where there are only a few possible values (e.g., sex).
- percentages of people for each value may be considered.

Population: is the collection of all items, units, measurements, or counts that are of interest in a particular study.

Examples

- All the cases with hepatitis B collected in a hospital from Jordan.
- All the deaths from COVID-19 in a city of Jordan.

Sample: It is a part of the **population** (subset). The sample is selected to represent the whole population. Therefore, the **sample** can be considered a representative part of the **population**.

Examples

- A group of 15 patients from all the cases with hepatitis B collected in a hospital from Jordan.
- A group of 10 deaths from all the deaths from COVID-19 in a city of Jordan.

Important Notation: Population size is N and Sample Size is n where $n \leq N$.

Example

An example of a sample would be the names of 125 employees being chosen out of an employees directory from King Abdullah University Hospital (KAUH) of 1500 employees in Jordan. In this case, the population is all 1500 employees ($N = 1500$), and the sample ($n = 125$).

Inferential Statistics

Consists of methods that use sample results to draw conclusions (inferences) in order to help make decisions or predictions about a population.

Note that: Inferring the characteristics of a population from a sample is the central concern of statistical inference. To accomplish this aim, we need to develop a probability model (probability distribution).

Statistical Studies

There are several different ways to classify statistical studies. We will explain two types of studies as follows:

1- Observational Study

In this study, the researcher merely observes what is happening or what has happened in the past and tries to draw conclusions based on these observations.

2- Experimental Study

In this study, the researcher manipulates one of variables and tries to determine how the manipulation influences other variables.

Statistical Packages

Data analyses can be carried out using a statistical package.

- A statistical package is a collection of statistical programs that describe data and perform various statistical tests on the data.
- A few statistical packages include SAS, SPSS, Stata, MINITAB, and Excel.

IMPORTANT

Students should realize that the computer and calculator merely give numerical answers and save the time and effort of doing calculations by hand. The student is still responsible for understanding and interpreting each statistical concept.

Final Step

Publication of manuscript

The final step in this study, after completing the data analysis, was to compile the results in a publishable manuscript.

Note that: Essential data obtained in the data analysis step is included in the final version of this manuscript for publishing.

The End