

“Bio-Statistics – Concept, Definition



CONTENT



- What is Bio-Statistics?
- Definition of Statistics
- Distinction between
 - Data, Information, Statistics
- Types of data
- Measurement scale of data
- Functions of Statistics
- Limitations of Statistics

What is Biostatistics ? (biometrics or biometry)

- ☞ Development and application of statistical techniques to scientific research relating to life:
 - ☞ **Human, Plant and Animal**
 - ☞ Here the focus is on Human life and Health. Thus the areas of application relates to:
 - **Pharmacology,**
 - **medicine,**
 - **epidemiology,**
 - **public health,**
 - **Physiology and anatomy,**
 - **Genetics**

Essence of Bio-Statistics?

***AS the knife is to Surgery
so is the Bio-satistics
to Medical Research***

What is Statistics? (Definition)

- The Statistics is defined differently by different authors from time to time.
- The reasons being:
 - Its scope and utility has widened considerably over time.
 - It is defined in two ways:
as “statistical data” and
“statistical methods”

DEFINITION OF STATISTICS IN DATA SENSE

Statistics is defined as the aggregate of facts

- *Effected to a marked extent by multiplicity of causes,*
- *numerically expressed,*
- *enumerated or estimated according to a reasonable standard of accuracy,*
- *Collected in a systematic manner for a predetermined purpose and*
- *placed in relation to each other.*

Prof. Horace Secrist

DEFINITION OF STATISTICS AS A BODY OF SCIENCE

American Heritage Dictionary® defines statistics as: "The mathematics of the Collection, organization, and interpretation of numerical data, especially the analysis of population characteristics by inference from sampling."

The Merriam-Webster's Collegiate Dictionary® definition is: "A branch of mathematics dealing with the collection, analysis, interpretation, and presentation of masses of numerical data."

A Simple but Concise definition by Croxton and Cowden:

“Statistics is defined as the Collection, Presentation, Analysis and Interpretation of numerical data.”

In the line of the definition of Croxton and Cowden, a comprehensive definition of Statistics can be:

“Statistics defined as the science of

- **Collection,**
- **Organisation,**
- **presentation,**
- **analysis and**
- **interpretation of numerical data.”**

COLLECTION OF DATA

This is the first and most important stage of statistical investigation.

Since the data are the raw material to any statistical investigation, utmost care is necessary to be taken for

collection of reliable and accurate data.

The first hand collection of data is very difficult. Therefore the investigator should see if the data intended for the purpose is already collected by some other agencies. This would help

*save time,
money and
duplication of effort.*

ORGANISATION OF DATA

The data collected from a published secondary source is in organised form. The large mass of data collected from a survey need to be organised.

The first stage in organising data is **editing**.

- Editing for completeness,
- Editing for inconsistencies,
- Editing for homogeneity,
- Editing for accuracy,
- Editing for reliability,

The 2nd stage in organising data is **classification of data**.

That means arranging data according to certain common characteristics of the units on which data are collected.

The final step is the **tabulation**.

The purpose is to present data in rows and columns so that there is absolute clarity in data to be presented

PRESENTATION OF DATA:

The presentation of data is an art and science. It facilitate
**statistical analysis,
comparison and
appreciation.**

The data can be presented through:
Diagrams, Graphs, Pictures etc.

ANALYSIS OF DATA:

Analysis of data are done through;
**Simple observation,
Application of simple and**

**highly sophisticated statistical techniques like;
Measures of central tendency,
variation, correlation, regression,
seasonality, trend, computation of indices etc.**

INTERPRETATION:

This refers to drawing conclusion from the data collected and analysed.

This requires high degree of skill and experience.

Wrong interpretations lead to fallacious conclusion and decisions.

Correct interpretation lead to valid decision and aid in taking decisions.

According to Prof. Ya Lun Chaou “ ***Statistics is a method of decision making in the face of uncertainty on the basis of numerical data and calculated risk***”.

DATA, INFORMATION AND STATISTICS

DATA

“Facts or figures from which conclusions can be drawn”.

- Data can take various forms, but are often numerical.
- Data can relate to an enormous variety of aspects, for example:
 - the daily weight measurements of each individual in your classroom;
 - Blood pressure and pulse recording of patients at regular interval

INFORMATION:

“Data that have been recorded, classified, organized, related, or interpreted within a framework so that meaning emerges”.

Information, can take various forms:

- The number of persons in a group in each weight category (20 to 25 kg, 26 to 30 kg, etc.);
- Number of Adolescents having menarch cycle within <20 days, 20-25 days, 25 – 30 days, etc.
- Number of patients responding to the drug within 3-5, 5- 7 & more than 7 days of therapeutic dose

- ***Just as trees are the raw material from which paper is produced, so too, can data be viewed as the raw material from which information is obtained.***

STATISTICS

☞ A statistics is "a type of information obtained through mathematical operations on numerical data".

STATISTICS EXAMPLE

- **the average weight of people in your classroom;**
- **the average number of in-door patients per day in a month for 12 months.**
- **the minimum and maximum blood pressure observed each hour of critical patient.**

Data collected on the weight of 20 individuals in your classroom

Data	Information	Statistics
20 kg, 25 kg 28 kg, 30 kg, ... etc.	5 individuals in the 20-to-25-kg range	Mean weight = 22.5 kg
	15 individuals in the 26-to-30-kg range	Median weight = 28 kg

TYPES OF DATA

- **Qualitative Data:** They are Nonnumeric.

In form of words such as

- Description of events: *Verbal Autopsy*
- Transcripts of interview: *Opinion of CDMOs on Efficacy of ASHA*
- Written documents: Guidelines on Roles and Responsibility of ASHA
- Characteristic not capable of Quantitative measurement: *Sex of a patient, Colour, Odour*

- **Quantitative Data:** They are numeric.

The characteristic capable of quantitative measurement: Height, Weight, Blood pressure etc.

MEASUREMENT AND MEASUREMENT SCALES

- Measurement: Assignment of numbers to object/observation or event according to a set of rules.
- Measurement Scale: Measurements carried out under different sets of rules results in different categories of measurement scale.
 - 4 categories of measurement scale
 - Nominal scale
 - Ordinal scale
 - Interval scale
 - Ratio scale

NOMINAL SCALE

- The lowest measurement scale.
- It consist of naming observations or classifying them into mutually exclusive and collectively exhaustive categories.
- The practice of using numbers to distinguish among the various medical diagnosis constitutes measurement on a nominal scale.
- It includes such dichotomies as male - female, well - sick, under 65 years of age - 65 and over, child - adult, and married-not married.

ORDINAL SCALE

- **This is an improved scale than nominal**
- **Whenever observations are not only different from category to category but can be ranked according to some criterion, they are said to be measured on an ordinal scale.**
- **In this scale data can be compared as less than or greater than.**
- **But the difference between two ordinal values can't be 'quantified'.**

ORDINAL SCALE

Examples:

- ❑ **The Nutritional Status of a group of Children Classified as Normal, Gr. 1, Gr. 2, Gr. 3, Gr. 4 level of malnutrition.**
- ❑ **Convalescing patients classified as Unimproved, Improved, Much improved.**
- ❑ **Socio-economic Status: Low, Medium and High**
- ❑ **Appropriate Statistics:**
 - ❑ **Non-parametric statistics are used for analysis of ordinal scale data.**

INTERVAL SCALE

- This is more sophisticated than the previous two
- In this scale it is possible to order measurements and also the distance between any two measurements can be quantified.
- There is a use of unit distance and zero point, both of which are arbitrary.
- The selected Zero is not necessarily a true Zero in that it does not have to indicate a total absence of the quantity being measured.
- It is truly quantitative scale

Example:

- **The way in which temperature is usually measured (degrees Fahrenheit or Celsius). The unit of measurement is the degree and the point of comparison is the arbitrarily chosen “Zero degrees,” which doesn’t indicate a lack of heat. The interval scale unlike the nominal and ordinal scales is a truly quantitative scale.**
- **Statistics used: Parametric statistical technique is used.**

RATIO SCALE

- It is the highest level of measurement scale
- You are also allowed to take ratios among ratio scaled variables.
 - Physical measurements of height,
 - weight,length are typically ratio variables.

It is now meaningful to say that 10 m is twice as long as 5 m. This ratio hold true regardless of which scale the object is being measured in (e.g. meters or yards). This is because there is a natural zero.

Statistics used: Parametric statistical technique is used.

Functions of Statistics

- It presents facts in a definite form
- It simplifies mass of figures
- It facilitate comparison
- It helps in formulating and testing hypothesis.
- It helps in prediction
- It helps in formulation of suitable policies.

Importance of Statistics

- **The Statistics is viewed not as a mere method of data collection, but as a means of developing sound techniques for their handling and analysis and interpretation and helping to take decision in the face of uncertainty.**
- **It provides crucial guidance in determining what information is reliable and which predictions can be trusted. They often help search for clues to the solution of a scientific mystery, and sometimes keep investigators from being misled by false impressions.**

Importance Statistics

Contd.

- **Therefore the importance of bio-statistics is increasing day by day and its scope is widening day by day.**
 - There is hardly any sphere of human life where Statistics is not being used.

Limitations of statistics?

- **However wonderful statistics are, they are not flawless and have some very important limitations.**
- **How representative is your sample?**
 - ❑ ***Samples used in statistical tests that do not represent the population adequately can give reliable results but with little relevance to the population that it came from.***

Limitations of statistics?

- ❑ Most of us use samples to represent a population, with the number of data points collected in the sample depending on the resources available. If we were to catch the nearest beetle and measure its length, would we be able to say that length is typical (average) of that species? Not with any degree of confidence! It may have just emerged as an adult, or it may have just escaped from a growth-hormone development laboratory and is 20 times the biggest size ever recorded before. For reasons such as these we must collect as many measurements as we can to get an idea of the variability within a population. Then we can find out how representative the sample is by calculating standard error. The lower the standard error is relative to the mean the closer the sample is to representing the population.

Limitations of Statistics

➤ How well does your data fit the requirements of the tests?

- ❑ All statistical tests are limited to particular types of data they can be applied to. For instance, some tests are based on the data having a normal distribution. By using this "normal" pattern in the data they provide us with the results.
- ❑ Results based on data with strong departures from the requirements of the test used will be less reliable than results from data that meet the requirements of the test.

Limitations of Statistics

- **How well the statistics is suited to the study of qualitative phenomena**
 - ❑ Statistics deals with phenomena capable of being expressed numerically.
 - ❑ The qualitative phenomena like honesty, integrity, poverty can not be analysed directly by statistical technique. However, they can be analysed once they are converted to some quantitative scores on suitable scales of measurement.
- **Statistics does not study individuals**

Statistics deals with aggregate of facts and does not give specific recognition to the individuals items of a series. Individual items taken separately does not constitute statistical data and meaningless for statistical enquiry.

Limitations of Statistics

- **Statistical Laws are not exact**
 - ❑ Unlike physical sciences, the statistical laws are only approximation and not exact. Statistics gives conclusion or decision in terms of probability not in terms of certainty. Statistical conclusions are not universally true. They are true on an average.
- **Statistics is liable to be misused**

Limitations of Statistics

➤ **Statistics is liable to be misused**

- ☐ **It must be used by experts. It is most dangerous in the hands of inexperienced.**
- ☐ **The use of statistical tools by inexperienced untrained persons might lead to very fallacious conclusion.**
- ☐ **Statistics do not bear the label on their face regarding their quality and hence can be misused with ulterior motives by people with vested interest.**

“Beauty of Statistics lies in its capability of handling a large mass of data, scientific analysis backed by sound reasoning and logic skills and a little common sense to figure out problems”

B. B. Nanda- 28.10.2007

When you can measure what you are speaking about and express it in numbers, you know something about it but when you can't measure it, when you can't express it in numbers your knowledge is a meager and unsatisfactory one.

Lord Kelvin

And Always Remember

**THERE IS STRENGTH IN
NUMBERS**

Exercises

- Define Statistics and Biostatistics.
- Define the word *measurement*.
- List, describe and compare the four measurement scales.
- For each of the following variables indicate whether it is quantitative or qualitative and specify the measurement scale that is employed when taking measurements on each:
 - a) Class standing of the members of this class relative to each other.
 - b) Admitting diagnosis of patients admitted to a mental health clinic.
 - c) Weights of babies born in a hospital during a year.
 - d) Gender of babies born in a hospital during a year.
 - e) Range of motion of elbow joints of students enrolled in a university health sciences curriculum.

- f) Under-arm temperature of day-old infants born in a hospital.
- For each of the following situations, answer question a through e:
 - a) What is the sample in the study?
 - b) What is the population?
 - c) What is the variable of interest?
 - d) How many measurements were used in calculating the reported results?
 - e) What measurement scales was used?

Situation A. A study of 300 households in a small southern town revealed that 20% had at least one school-age child present .

Situation B. A study of 250 patients admitted to a hospital during the past year revealed that, on the average, the patients lived 15 miles from the hospitals.

Reference

- Biostatistics – A foundation for Analysis in the Health Sciences: Wayne W. Daniel, Seventh Edition, Wiley Students Analysis.
- A First Course in Statistics with Application: A.K. P. C Swain, Kalyani Publishers

The background is a solid blue color with a subtle grid pattern of thin, lighter blue lines. The grid consists of both horizontal and vertical lines, creating a series of small squares across the entire image.

THANK YOU