Biostatistics is used when tools of statistics are applied to the data that is derived from life science.

Characteristics of Bio-Statistics

(1) Statistics is the aggregate of facts.
(2) Statistics is numerically expressed.
(3) Statistics is usually affected by multiplicity of causes and not by single cause.
(4) Statistics must be related to some field of inquiry.
(5) Statistics should be capable of being related to each other, so that some cause and effect relationship can be established.
(6) The reasonable standard of accuracy should be maintained in statistics.

Importance of Bio-statistics

(1) Statistics help in presenting large quantity of data in a simple and grouped form.
(2) It gives the methods of comparison of data.
(3) In enlarges individual mind.
(4) It helps in finding the conditions of relationship between the variables.
(5) It proves useful in almost every sphere of human activities.

Application and Uses of Biostatistics

Biostatistics is applied and used in different branches of bioscience.

(I) In Physiology and Anatomy

(1) To define what is normal or healthy in a population and to find limits of normality in variables.
(2) To find the difference between the mean and proportion of normal at two places or in different periods.
(3) To find out correlation between two variables X and Y such as height and weight.

(II) In Pharmacology

(1) To know that action of drug-a drug given to animals and humans to observe the changes produced are due to the drug or by chance.
(2) To compare that action of two different drugs or two successive dosages of the same drug.
(3) To find out the relative potency of a new drug with respect to a standard drug.

(III) In Medicine

(1) To compare the efficacy of a particular drug. For this, the percentage of cured and died in the experiment and control groups is done.
(2) To find out an association between two attributes such as cancer and smoking.
(3) To identify signs and symptoms of a disease of syndrome. Cough and typhoid is found by chance and fever is found in almost every case.

(IV) In community medicine and Public health

(1) To test usefulness of sear and vaccines in the field-the percentage of attacks or deaths among the vaccinated subjects is compared with that among the unvaccinated ones to find whether the difference observed is statistically significant.
(2) In epidemiological studies – the role of causative factors is statistically tested.
(3) In public health, the measures adopted are evaluated.

(V) In Genetics

Biostatistics is used in studying the genetics. Mendel's laws of inheritance are tested by x2 test. Hardy Weinberg law is tested by bio statistical methods.

Scope of Biostatistics

Use of statistical methods are constantly increasing in biological sciences. The development of biological theories is closely associated with statistical methods.
Heredity, one of the recent branches of biology is mainly based on biostatistics.
Therefore, for the students of biology, the knowledge of biostatistics is a must.

Statistical Terms

Basic statistical terms: Population, Sample, Data. Observation. Parameter and Statistic. Characteristic. Attribute/Variable. Statistical error. Subscript and summations. Functions of statistics. Array. Class interval, Class size, Class mark. Freq. distribution

Term is a word used to explain a particular identity.
Symbol is a mark or sign with a particular meaning. The use of terms and symbols allow statisticians to deal with general expression and general results. By doing so, considerable saving in space and time takes place.

Population may be defined as "an entire group of organisms of one species, occupying a definite area or study elements-persons, things or measurements having some common fundamental characteristics". in statistics population is a well-defined group which is being studied. Suppose one has to study the incidence of helminth infection in rabbits. For this purpose, 100 rabbits are collected randomly and brought in the laboratory. Here 100 rabbits are population for this experiment and result will represent the universal population of rabbit. In other words, in statistics, population always means the total number of individual of individual observations from which inferences are to be made at a particular time.

Sample-The selected part of a population is known as sample.
For example, all patients of AIDS of the world represents a population, whereas individual observations on 10 or 20 or 30 (any convenient number) patients from the population refer to a sample.

Data:
Data is a collection of observations expressed in numerical figures. Data is always in collective sense and never be used singular. The data in statistics are generally based on individual observations. The Hb% of 10 patients suffering from Kalaazar was measured as 10.2, 9.6, 8.8, 10.7, 9.9, 10.8, 11.3, 9.5, 8.9, 8.8, mg/100 ml. Here 10.2, 9.6.....8.8 mg/100 ml. are a set of values for an event i.e. Hb% and is called data.

Primary data
The data which are collected directly by an investigator for the first time for a specific purpose are called as **primary data.** These are raw data or data in original nature, and directly collected from population. The collection of primary data may be made through either by complete enumeration or sampling survey methods.

Secondary data
If data are collected and used by any other agency at least once then such data are termed as secondary data.
Note: In scientific researches only primary data are used.

Qualitative and quantitative data

Qualitative: According to quality attributes the data is called qualitative.
For example, lions of **Gir Sanctuary** of Gujarat State are to be classified in respect to one attribute say sex, in two groups, one is of **male** and the other is of **female.**

Quantitative: According to magnitude the data is called quantitative.
For example, chickens of a poultry farm may be classified on the basis of their growth rate.
Quantitative data may also be classified into two types:
(i) Continuous. Values of variate **do not exhibit any breaks or jumps.**
For example, the increasing length and weight of a child.
(ii) Discrete. Values of variate **vary by infinite jumps.**
For example, the oxygen consumption of rat **(*Rattus rattus*)** of different weight groups were measured as 500 cc/h/100 ml, 600cc/h/100 ml, 620 cc/h/100 ml, 680 cc/h/100 ml and so on.

Observation
Measurement of an event is called observation.
For instance, blood pressure, temperature of body, oxygen consumption etc. are events whereas, 160 mm and 80 mm. (upper and lower pressure), 106oF, 65 kg/hour/100 ml are their respective observations. The source that gives observations such as object, person etc. are called observational units. In biostatistics statistics the term individuals or subjects is used for observational units.

Parameter and Statistics
A value calculated from a defined population such as
Mean (μ for population)
Standard deviation (σ)
Standard error of difference of mean a (X_1- X 2) etc.
are called a **parameter.** It is a constant value because it covers all members of the population. Familiar examples are mean height, birth rate (fecundity) and mortality rate etc. of any one species of animals or plants. The quantity calculated to represent a **character of population** is known as **parameter** whereas quantity calculated to present the **character of the sample** is called statistics. The former is the constant quantity whereas the latter is variable. In other words, we can say that the numerical quantities which characterise a population (in respect of any variable) are called parameter. For example; if the variable is in length and the measurement of length is taken for the entire population, the mean length can be regarded as parameter. But the mean length of the sample (X) can be regarded as statistic. Standard deviation of sample (s) and Standard error of difference of means (X 1 - X 2)] of sample are statistic. In biological experiments values calculated from a large sample is often applied to population and may be a valid estimate of population. Therefore, in biostatistics, though not desirable, parameter and statistics are often used as synonyms.

Characteristic
The term 'characteristic' means a quality possessed by an individual i.e., object, item of population. Height, weight, age, Hb%, VO_2 etc. are characteristics.
In statistics, characteristics are of two types.
(1) Non-measurable 'characteristics' is called **Attributes.**
(2) Measurable 'characteristics' is called **Variables.**

Attributes: Attributes are the non-measurable characteristics which not be numerically expressed in terms of unit. These are qualitative object. For example: sex, illiteracy etc.

Variables. Variables are the measurable characteristics which can be numerically expressed in terms of some unit. These are quantities which are capable of being measured by quantitative methods directly. An individual observation of any variable is known as variate. If we measure the height of some individuals of a population and obtain some values, the obtained values is variable. For example, height and length in cm, weight in g, Hb in %, oxygen consumption in VO_2/100 ml etc. of individuals.

A variable is a symbol, such as *X, Y, Z* etc., that may take any value in some specified set of numbers. Whatever the value of the variable actually observed is the actual value and is denoted by the $X_1, X_2, X_{3\ldots\ldots}, X_n$ or $Y_1, Y_2, Y_3, \ldots\ldots, Y_n$ or $Z_1, Z_2, Z_3, \ldots\ldots, Z_n$. The possible values of a variable are those values that the variable may possibly take. For example, the "weight of fishes in a pond" may take any value between $X_1, \ldots\ldots X_n$ *g*. Both X_1 and X_n are inclusive.
Depending on the break or continuity, variables are of two types:
(a) Discrete variable is one which cannot take all the values and there is a gap between one value and the other. For example, the number of zooplanktons in 5 ponds were obtained as 98, 305, 387, 105 and 208. So, in this case number of zooplankton is the discrete variable. Number of persons in a family, no. of books in. a library is discrete variable, because they cannot take functional value. One cannot say that there are 3.5 persons in my family or there are 500.6 books in a library. The discrete variable may take any integer value from 0 to ∞.

(b) ***Continuous variable is*** the one which can take any values and there is no interval. For example, the weight and height of human being is a continuous variable because it may take any value. Height of students in a class may be 120 cm, 120.2 cm, 120.5 cm, 120.7 cm, 120.9 cm and so on. Measurement of Hb%, VO_2 consumption etc. present continuous variable. Generally speaking, 2 discrete variable take integer value while continuous variables take fractional values.

Statistical error
In statistical terminology, the word 'error' is used in special sense. Error shows the extent to which the observed value of a quantity exceeds the true value.
Error =Observed value - True value.