

Definition of Biostatistics/ Biometry

According to Horace Secrist

‘By Statistics we mean aggregates of facts affected to a marked extent by multiplicity of causes numerically expressed, enumerated or estimated according to reasonable standards of accuracy, collected in a systematic manner for a pre-determined purpose and placed in relation to each other.’

Meaning of Biometry

Statistics is a branch of science which deals with methods of collection, classification and analysis, i.e., drawing inferences from collected data; then testing of hypothesis and making comment on these.

The statistics can be applied in all scientific branches like social, physical and biological. The use of statistics in Biology is known as Biostatistics or Biometry. It deals with the application of different statistical methods and analysis of data collected from biological system.

Characteristics of Biostatistics/ Biometry

Statistics should contain the following characteristics:

- i. Numerical data should be aggregates of facts.
- ii. Data should be affected to a marked extent by multiplicity of causes.
- iii. Data should be numerically expressed.
- iv. Data should be enumerated or estimated according to reasonable standards of accuracy.
- v. Data should be collected in a systematic manner.
- vi. Data should be collected for a predetermined purpose.
- vii. Data should be placed in relation to each other.

Classes of Statistics

Webster defined statistics in the same sense

‘Statistics are the classified facts respecting the conditions of the people in a state ... specially those facts which can be stated in numbers or in tables of numbers or in any tabular or classified arrangement.’

According to Seligman:

‘Statistics is the science which deals with the methods of collecting, classifying, presenting, comparing and interpreting numerical data collected to throw some light on any sphere of enquiry.’

King defined statistics as ‘the science of statistics is the method of judging collective, natural or social phenomenon from the ‘results obtained from the analysis or enumeration or collection of estimates.’

A. L. Bowley suggested statistics variously as

- i. Science of counting,
- ii. Science of averages,
- iii. Science of measurement of the social organism regarded as a whole in all its manifestations.

Boddington defined statistics as the science of ‘estimates and probabilities’, while, Lovitt* defined it as collection, classification and tabulation of numerical facts as the basis for explanation, description and comparison of phenomena.

Thus, statistics may be defined as:

- i. Collection,
- ii. Presentation,
- iii. Mathematical analysis,
- iv. Logical analysis of the numerical data.

Statistics has 2 main classes:

Statistical methods are devices, rules of procedure and general principles which are applicable to all kinds or groups of data. Thus, they include all the general principles and techniques which are commonly used in the collection, analysis and interpretation of data relating to any sphere of enquiry.

Applied statistics deals with the application of statistical methods to specific problems or concrete forms.

When the tools (principles) of statistics are applied to study living organisms or in other words it can be said that numerical data obtained from biological organisms it is called Biostatistics or Biometry.

Variables:

Data can be of 2 types, such as:

Qualitative, e.g., flower colour, leaf shape, pupil colour of eye, etc., these are non-numerical characters.

Quantitative, e.g., plant height, number of branches per plant, seeds per capsule, fish yield, etc. The item representing the quantitative characters is known as variable, that is, a character differing in value from one organism to another, from one sample to another.

Again variable may be discrete or continuous.

Discrete the units are obtained by counting and such variables are always represented as integers, e.g.,

<i>Plant No.</i>	<i>No. of capsules/plant</i>
1	17
2	25
3	20
4	8
5	15
6	17

Continuous variable which can take into account any value within a certain range, e.g.,

<i>Plant No.</i>	<i>Soluble seed protein content (gm/gm of tissue)</i>
1	0.227
2	0.281
3	0.196
4	0.248
5	0.256

Population and Sample:

If one had to count the number of Rhododendron plants in the hills of Darjeeling then it is practically impossible to do it. In such case, smaller pockets

are taken which bears characteristics of the Darjeeling hills and subsequently survey can be made. Thus, Darjeeling hills is the population and smaller pockets taken are the samples. So, samples are true representative of the population.

Random Sampling:

Collection of data unbiasedly. Selection is done purely on chance occurrence only. Size of the sample is very essential for accuracy. Bigger size sample gives more accuracy; however, unmanageable size is often unnecessary.

Statistical Error:

It is difference between the true value and the estimated value.

Errors may arise due to instability of the collected data. Such errors are called errors of origin. Errors may also arise on account of manipulation in counting, measurement, description or approximation. Such errors are known as errors of manipulation. However, errors arising due to inadequate size of the sample are known as errors of inadequacy.

Errors may be biased or unbiased and can be measured either absolutely (difference between true value and estimated value) or relatively (ratio of the absolute error to the estimate).

Application of Biometry

In biological system, it is very difficult to draw any concrete conclusion about any happening, as there is much difference among the individuals. Such as in the same plant, the leaves are all different; again in a field, all the same kind of plants are not identical; in the same place, the same crop does not respond equally in different plots. So it is very difficult to reach at a definite and reliable conclusion unless we take the help of statistical analysis of the observed data.

In case of plant breeding, the application of biometry has immense importance, as the statistical analysis only helps to conclude about a population of plant; such as:

(a) How the character, which is acquired by the population, is distributed, how the character is being inherited, etc.

(b) Whether two characters in a population are correlated or not, if correlated, how much they are correlated.:

- (c) In case of plant breeding whether Mendelian Genetics is being followed or not, can be tested by chi-square test.
- (d) The application of fertilizer or irrigation or any kind of cultural practice can be effectively justified by application of test of significance.
- (e) The laws of probability is of importance in genetics for forecasting the chance of obtaining certain result from a cross and elucidating the operation of genetic principle.
- (f) The effects of two or more different fertilizer application on the same field can be analysed by analysis of variance.
- (g) Study of alleles of genes in populations and the factors which maintain or change the frequencies of particular genotype in populations can be studied.

Limitation of Biometry:

- (a) All the conclusions about statistical analysis depend on the availability of sample data. If the sampling is biased, the analysis will be erroneous.
- (b) Statistics can be used and applied only on collective data, not on individual data.
- (c) Statistical laws are always true in case of large population on the average, not on the small sample.