

SUBJECT: ELECTIVE PAPER-2
BIOSTATISTICS AND SEEDTECHNOLOGY

SUBJECT CODE: 18MBO24E

UNIT - 1

Dr. P. Latha

Assistant Professor
PG and Research Department of Botany
Government Arts College (autonomous)
Coimbatore -18
Mobile No : 9894275795

1 Introduction to Biostatistics

Biostatistics is the science of collection, analysis and interpretation of facts and numbers connected with Biology.

Biostatistics is also called *Biometrics*. Biometrics refers to *biological measurements*.

A simple example to explain Biostatistics is the estimation of oxygen in a few water samples. The water samples form a *population*. The estimation of O_2 in each water sample is the *collection of data*.

The amount of oxygen in the water samples is the *data*.

Arranging values in columns is called *tabulation*.

In one water sample the amount of oxygen will be higher and in another it is lower. It is *interpretation*.

Table 1.1: A simple table to explain the science of Biostatistics.

Water samples	Amount of O_2 in ml
1	4.5
2	6.9
3	6.2
4	5.3

2

Scope of Biostatistics

Without a proper statistical data, it is not at all possible to draw a conclusion for most biological problems. Hence, statistics is considered to be an *important tool* in biological investigations. In Biology, statistics is required for :

- Presenting facts in a definite order.
- Simplifying the mass of figures.
- Facilitating easy *comparison*.
- Formulating and *testing hypotheses*.
- Predication of *correlations* between cause and effect.
- Formulation of suitable *biological indices*.
- Developing new policies.

4 Data

The values recorded in an experiment or observation are called **data**. The word 'data' can be used both as singular and plural. There is no word like 'data'.

The data is of two types, namely **primary data** and **secondary data**.

The data collected by an investigator is called **primary data**. It is the first hand information. The person collecting the data is called **investigator**.

The data collected from another source is called **secondary data**.
Eg. Data collected from News papers, journals, etc.

Primary Data

The data collected for the first time is called **primary data**.

Primary data gives the **first hand** information. It is **original** in nature. It is **accurate** and **reliable**.

It consumes more money and time.

The amount of O_2 estimated by a student by titration is a primary data. The rainfall recorded by PWD department is a primary data.

Secondary Data

The data obtained from already collected data is called **secondary data**.

They are **secondary** in nature. They are economical. They are **not reliable**.

The amount of O_2 estimated by a student by titration is a primary data. When this data is used by a teacher, it becomes a secondary data.

The rain fall data obtained from a PWD department is a secondary data.

Table 4.1: Distinction between primary and secondary data.

Basis	Primary data	Secondary data
1.Originality	It is original, because the investigator himself collects the data.	It is not original. The investigator makes use of the data collected by other agencies.

6

Methods of Data Collection

The main methods of collecting the data are as follows:

1. *Direct personal investigation*
2. *Indirect oral investigation*
3. *Information through correspondents*
4. *The questionnaire method*
5. *Experiments*
6. *Census*
7. *Sampling.*

1. Direct Personal Investigation

The investigator has to contact the sources of information directly and personally. This method of collecting data is suitable: a. When the area of investigation is limited, b. When a higher degree of accuracy is needed, c. When the results of the investigation are to be kept confidential, d. Where the area of investigations is complex and heterogeneous and e. Where a greater degree of intellectual power, experience and labour are required.

2. Indirect Oral Investigation

There are certain situations when the informations cannot be approached directly. In such situations, we have to collect data from those persons who may possess some knowledge about the investigation. Such persons are known as '*witnesses*'.

3. Information Through Correspondents

In this method, the investigator does not collect the information directly from the respondents. The task is handed over to a few persons or a group of persons, who are known as *correspondents*.

4. Questionnaire Method

In this method, the information is obtained from the informants through questionnaire.

5. Experiments

Data can be collected by doing experiments.

6. Census

Census is a method of *collection of data*. Census is the *counting* of all the members of a population one by one.

Eg. *The trees in a coconut grove are counted by census method.*
Human population is assessed by census.

The results obtained in census are reliable and accurate. Data is obtained from each and every individual.

It is very expensive. It consumes time and labour.

7. Sampling

Sampling is a method of *collection of data*.

Sampling is the process of getting a representative fraction of a population.

Sample is the representative fraction of a population.



14 Sampling

Sampling is a method of *collection of data*.

Sampling is the process of getting a representative fraction of a population.

Sample is the representative fraction of a population.

In sampling method, a small group is taken from a large population. This small group is the **sample**. Analysis of the sample gives an **idea** of the population.

When the population is very large or infinite, sampling is the suitable method of data collection.

One rice is tested from a pot of boiling rice to arrive at a conclusion.

In an electric bulb factory, the bulbs are tested at intervals how long they will burn. If all are tested there is nothing left for selling.

One grape is tasted before buying a bunch of grapes.

The oxygen content of pond water can be found out by titrating just 100 ml of water.

The length of leaves of a neem tree can be calculated by measuring just 10 leaves.

The food habits of all students of a college can be understood by observing only about 100 students.

A small representative fraction of a population is called a sample.

Getting a sample from a population is called sampling.

Sample, refers to a part of the population selected for analysis.

The part is the representative of the whole population. For example, we want to study the average height of the students studying in the Higher Secondary schools in Palayamkottai. It is not necessary to observe the height measurements of all the students. What we can, in fact, do is that we can take a small representative sample of a few students from different schools; take their height measurements and the result. Likewise we

want to know about the pattern of consumption expenditure of residents in certain locality. We can select a sample of households and collect the relevant data from them. The emerging pattern will be representative of the whole sample.

Advantages of Sampling

1. This method is comparatively more economical; it saves time, expenditure and energy because of the limited field of investigation.
2. In this method, a higher degree of accuracy is possible.
3. This method is organizationally more convenient as it does not require more number of investigators.
4. It is more scientific and hence reliable results can be obtained.
5. It is most suited to those places and situations where census method cannot be applied.
6. If '*approximate*' or '*aggregate*' results are to be known this type of enquiry is useful.

Disadvantages of Sampling

1. This method is not suitable if one is interested in the characteristics of individual constituents of the population.
2. In this method, a careful, proper plan is needed, otherwise incorrect, biased and misleading results will be obtained.
3. In this method, the services of experts are needed and if there is paucity of such personnel, it will yield unsatisfactory results.
4. In this method, the sources of error are larger.
5. Selection of an appropriate method of sampling is also a big problem. If the sample is not truly representative of the population, results will not be appropriate.

21

Frequency Distribution

Frequency distribution is a statistical table containing groups of values according to the number of times a value occurs.

Frequency means the number of times a particular value occurs. It is the repetition of values.

In frequency distribution, the data are arranged in an order in a table.

In a *frequency table* the data are grouped into *classes*. The number of values in each class is called a *frequency*.

There are 3 types of frequency distribution, namely,

1. *Frequency distribution without class intervals.*
2. *Frequency distribution with class intervals.*
3. *Cumulative frequency distribution.*

Problem - 5

The age at marriage of 25 men and women is given below. Arrange the data in the form of a frequency distribution.

Table 21.17

Age of Man	Age of Women	Age of Man	Age of Women
24	17	27	19
26	18	25	19
27	19	27	20
25	17	26	19
28	20	25	17
24	18	26	20
27	18	26	17
28	19	26	18
25	18	24	17
26	19	25	18
25	17	26	19
26	18	27	20
		24	29

Solution

It is an example of a bivariate distribution. The data in its arranged form, can be presented as follows:

Table 21.18

		Age of Man					
		24	25	26	27	28	F
Age of Woman	17	11	111	1			6
	18	1	11	111	11		8
	19	1	1	111	1	1	7
	20			1	11	1	4
f		4	6	8	5	2	25