

2- Analysing Biological Data

Biostatistics involves applying statistical techniques to analyse data. This analysis can uncover trends, correlations, and patterns. For example, analysing the growth rates of plants under different environmental conditions can reveal how factors like light and water affect growth.

3- Interpreting Results

After analysing data, biostatistics helps to interpret the results in a meaningful way. For example, interpreting the results of a survey on the prevalence of a disease in a population can guide public health interventions.

4- Predicting Outcomes

Biostatistics can be used to create models that predict future outcomes based on current data. For instance, predicting the spread of an infectious disease within a population helps in planning vaccination campaigns and allocating resources.

5- Public Health and Policy Making

In public health, biostatistics provides evidence-based insights that guide policy decisions and health guidelines. For example, statistical analysis of data on COVID-19 rates can lead to the implementation of COVID-19 vaccination campaigns.

Examples of the Uses of Biostatistics

1- Epidemiology

Epidemiologists use biostatistics to study the distribution and determinants of health and diseases in populations. For example, analysing data on COVID-19 infection rates, recovery rates, and the effectiveness of vaccines involves biostatistical methods.

2- Genetics

Biostatistics is used in genetic research to analyse the inheritance of traits and the association of genetic variations with diseases. For instance, genetic studies use biostatistics to identify genetic markers linked to diseases like diabetes and cancer.

3- Agriculture

In agricultural research, biostatistics helps in analysing crop yields, the effectiveness of fertilizers, and the resistance of plants to pests and diseases. For example, comparing the yield of different wheat varieties under various farming practices involves statistical analysis.

4- Clinical Trials

Biostatistics is crucial in the design and analysis of clinical trials that test new treatments and drugs. For instance, determining whether a new medication is more effective than a placebo requires rigorous statistical testing to ensure the results are statistically significant.

11.2- MEAN, MEDIAN, AND MODE

The mean, median, and mode are the measures that help summarize and understand data sets. The mean provides an overall average, the median gives the middle value, and the mode highlights the most frequent value.

Mean

The mean, also known as the average, is the sum of all the values in a data set divided by the number of values. It represents the central value of a data set.

Formula

$$\text{Mean} = \frac{\text{Sum of All Data Points}}{\text{Number of Data Points}}$$

Example

Consider the following data set representing the heights (in cm) of five students:

150, 160, 165, 155, 170.

$$\text{Mean} = \frac{150 + 160 + 165 + 155 + 170}{5} = \frac{800}{5} = 160$$

So, the mean height is 160 cm.

2- Median

The median is the middle value of a data set when the values are arranged in ascending or descending order. If the number of values is odd, the median is the middle value. If the number of values is even, the median is the average of the two middle values.

Steps to Calculate Median

1. Arrange the data in ascending order.
2. If the number of values (n) is odd, the median is the value at the position $\frac{n+1}{2}$.
3. If the number of values (n) is even, the median is the average of the values at positions $\left(\frac{n}{2}\right)$ and $\left(\frac{n}{2} + 1\right)$.

Example

Consider the data set: 150, 160, 165, 155, 170.

1. Arrange in ascending order: 150, 155, 160, 165, 170.
2. Number of values (n) = 5 (odd)
3. Median is the value at position $\left(\frac{5+1}{2}\right) = 3$.

So, the median height is 160 cm.

For an even number of values, consider the data set: 150, 160, 165, 155.

1. Arrange in ascending order: 150, 155, 160, 165.
2. Number of values (n) = 4 (even).
3. Median is the average of the values at positions $\left(\frac{4}{2}\right) = 2$ and $\left(\frac{4}{2} + 1\right) = 3$.

$$\text{Median} = \frac{155+160}{2} = \frac{315}{2} = 157.5$$

So, the median height is 157.5 cm.

3- Mode

The mode is the value that appears most frequently in a data set. A data set may have one mode, more than one mode, or no mode at all.

Steps to Calculate Mode

1. Count the frequency of each value in the data set.
2. The value with the highest frequency is the mode.

Example 1

Consider the data set: 150, 160, 165, 155, 160.

- Frequencies: 150 (1), 160 (2), 165 (1), 155 (1).
- The value with the highest frequency is 160.

So, the mode of the data set is 160.

Example 2

Consider the data set: 150, 160, 160, 155, 155.

- Frequencies: 150 (1), 160 (2), 155 (2).
- The values with the highest frequency are 160 and 155.

So, the data set is bimodal with modes 160 and 155.

Example 3

Consider the data set: 150, 160, 165, 155, 170.

- Frequencies: 150 (1), 160 (1), 165 (1), 155 (1), 170 (1).
- All values have the same frequency.

So, this data set has no mode.

11.3- BAR CHART

A bar chart is a graphical representation of data using bars of different heights or lengths. It is used to compare the quantities of different categories. Bar charts are effective for comparing different categories and visually representing the distribution of data.

Steps to Create a Bar Chart:

1. Gather the data to be represented in the bar chart.
2. Arrange the data into categories and their corresponding values.
3. Draw a horizontal axis (x-axis) and a vertical axis (y-axis).
4. Label the x-axis with the categories and the y-axis with the values.

- Determine the scale for the y-axis based on the range of values in the data set. Divide the axis into equal intervals.
- For each category, draw a bar with a height corresponding to its value. Ensure the bars are of equal width and are spaced evenly.
- Label each bar with its category name and value.

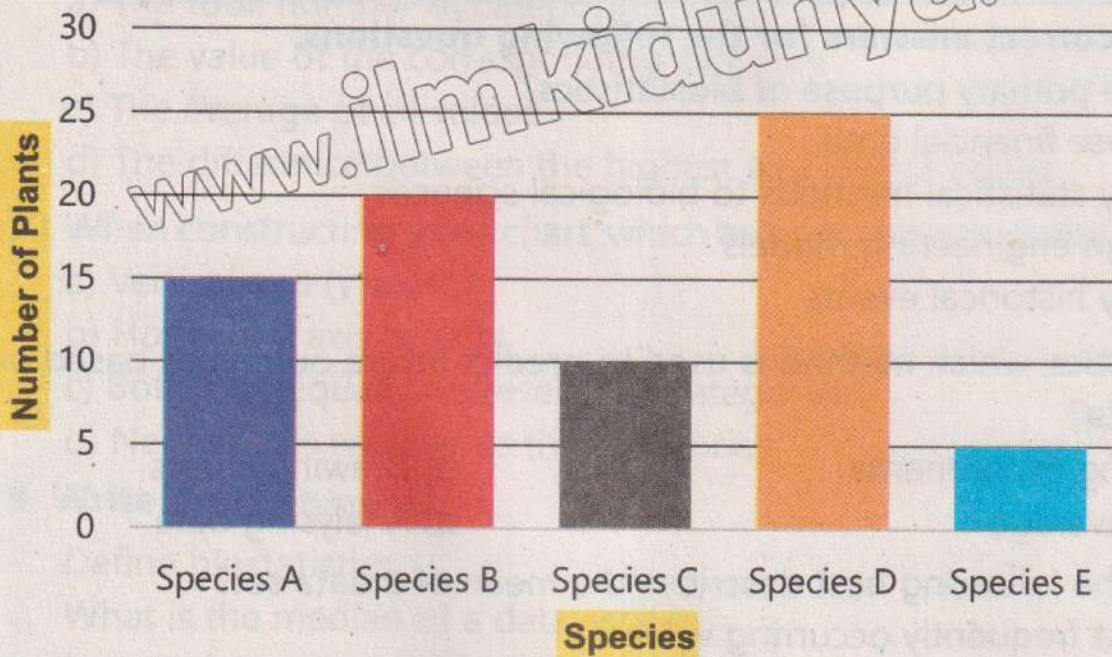
Example

Consider the following data representing the number of plants belonging to different species found in a field survey:

Species	Number of Plants
Species A	15
Species B	20
Species C	10
Species D	25
Species E	5

- Collect Data:** The data is already collected in the table above.
- Organize Data:** Data is organized in the table with species and their corresponding number of plants.
- Draw Axes:** Draw the x-axis and y-axis.
- Label the x-axis with the species:** A, B, C, D, E. Label the y-axis with the values i.e., number of plants.
- Scale the Axes:** The highest value is 25. Use a scale with intervals of 5 i.e., 0, 5, 10, 15, 20, 25.
- Draw Bars:** For each species, draw a bar up to the corresponding number of plants.
- Label the Bars:** Label each bar with the species name and its value.

Chart: Number of Plants



KEY POINTS

- Biostatistics is the application of statistical methods to biological sciences.
- Biostatistics helps in designing experiments, analysing data, interpreting results, predicting outcomes, and informing public health policy.
- Mean is the sum of all values divided by the number of values.
- Means provides an overall average, useful for understanding general trends.
- Median is the middle value when data is ordered. If even number of values, the median is the average of the two middle values.
- Median is useful for understanding the middle value, especially with skewed data.
- Mode is the value that appears most frequently.
- Mode highlights the most common value, useful for categorical data analysis.
- Bar charts help compare different categories.

EXERCISE

A. Select the correct answers for the following questions.

1. What is the primary purpose of biostatistics?
 - a) To analyse financial data
 - b) To apply statistical methods to biological sciences
 - c) To design engineering models
 - d) To study historical events
2. In biostatistics, which method is used to predict future outcomes based on current data?
 - a) Designing experiments
 - b) Drawing charts
 - c) Taking average
 - d) Analysing data
3. Which of the following best describes the mean of a data set?
 - a) The most frequently occurring value
 - b) The middle value when data is ordered
 - c) The sum of all values divided by the number of values
 - d) The difference between the highest and lowest values
4. If the data set is 5, 8, 12, 15, 20, what is the median?
 - a) 8
 - b) 12
 - c) 15
 - d) 20
5. What is the mean of the data set: 7, 8, 9, 10, 11?
 - a) 7
 - b) 8
 - c) 9
 - d) 10
6. When the number of values in a data set is even, how is the median calculated?
 - a) By choosing the middle value
 - b) By taking the average of the two middle values
 - c) By selecting the most frequent value
 - d) By adding all values and dividing by the total number of values
7. In a data set with values 3, 3, 6, 7, 8, 9, 9, what is the mode?
 - a) 3
 - b) 6
 - c) 7
 - d) Both 3 and 9
8. If a data set has no repeated values, what is the mode?
 - a) The highest value
 - b) The average of the data set
 - c) There is no mode
 - d) The median value

9. In a bar chart, what does the height or length of each bar represent?
- The total number of categories
 - The value of the corresponding category
 - The average of all values
 - The difference between the highest and lowest values
10. When constructing a bar chart, which axis usually represents the categories?
- Vertical axis (y-axis)
 - Horizontal axis (x-axis)
 - Both axes equally represent the categories
 - Neither axis represents the categories

B. Write short answers.

- Define biostatistics.
- What is the median of a data set?
- How is the mean calculated?
- What does the height of a bar in a bar chart represent?
- What is the mode of a data set?

C. Write answers in detail.

- Explain the importance of biostatistics in the field of public health. Provide examples of how it is used to inform public health decisions.
- Discuss the differences between mean, median, and mode. Include examples where each measure is most appropriate to use.
- Describe the steps involved in creating a bar chart using Excel. Include a discussion on how to customize the chart for better visualization and interpretation of data.
- Provide a detailed example of how to calculate the mean, median, and mode of a data set. Use the following data set for your calculations: 12, 15, 22, 8, 19, 25, 15.
- You are given the following data set, create a bar chart to represent the number of different types of fruits sold at a market in one week:
 - Apples: 30
 - Bananas: 45
 - Oranges: 20
 - Grapes: 25
 - Mangoes: 15

Ensure to label the axes and provide a title for the chart.