

(I) What is difference between Descriptive and Inferential Statistics?

→ Descriptive Statistics

Inferential Statistics

→ Descriptive statistics summarize and describe the main features of dataset. Inferential statistics use sample data to make conclusion & prediction about larger population.

→ It helps in organizing, analyzing and to present data and make hypothesis data in meaningful manner and predictions.

→ It deals only with the data that is collected or observed.

→ Common methods include mean, median, mode, standard deviation, etc.

→ Common methods include hypothesis testing, Confidence Intervals, correlation and regression analysis, etc.

→ It works with complete or actual data from a Sample or population.

→ It works with sample data to estimate or predict population parameters.

Ex: Calculating the average height of students in one class.

Ex: Predicting the average height of all students in the school using sample of classes.

(2) Differentiate between Discrete and Continuous Probability Distributions.

→ Discrete - Probability Distributions.

Continuous - Probability Distribution.

→ It deals with random variables that take countable outcomes.

→ It deals with random variable that take infinite outcomes within a range.

→ It can take only specific or separate values within a finite set. → It can take any value within a given interval.

→ It is represented using a probability mass function. → It is represented using a probability density function.

→ Each value has a specific probability greater than 0.

→ Probability of any exact value is 0, but there are found over intervals.

→ It displayed using bar graph, etc.

→ It displayed using a smooth curve.

Ex :- Number of students in class, number of calls received.

Ex :- Height, weight, temperature, time.

(3) What is importance of standard deviation and Variance in data analysis?

- Variance and standard deviation are statistical measures that describes how spread out the data around mean.
- Variance Measure the average of the square differences from the mean.
- Standard deviation is a square root of Variance showing spread in the same units.
- They are used in almost every field from analyze student performance and measuring product quality to assessing risk in investments.
- It helps to identify the consistency of data like low standard deviation means data points are close to mean.
- It helps to detect outliers and unusual data.
- In finance and machine learning it's used to measure risk and model accuracy.
- Ex :- two students score average at 80 marks.
A's scores are 78, 79, 83 It's smaller spread
B's scores are 60, 80, 100 It's larger spread.
- Both have same mean but student B has a higher standard deviation showing greater variation.

(4) Define the following,

- (i) Skewness
- (ii) Kurtosis
- (iii) Outlier.

→ (i) Skewness :-

→ It measures the asymmetry of a distribution indicating if data are skewed to the left or right of the mean.

→ (ii) Kurtosis :-

→ It measures that indicates how heavy or light the tails of distribution are compared to normal distribution.

→ (iii) Outlier :-

→ It is an observation that lies far away from the other data points in dataset.