

## 1. Differentiate between inferential statistics and descriptive statistics?

Ans:

BASIS FOR COMPARISON	DESCRIPTIVE STATISTICS	INFERENTIAL STATISTICS
Meaning	Descriptive Statistics is that branch of statistics which is concerned with describing the population under study.	Inferential Statistics is a type of statistics, that focuses on drawing conclusions about the population, on the basis of sample analysis and observation.
What it does?	Organize, analyze and present data in a meaningful way.	Compares, test and predicts data.
Form of final Result	Charts, Graphs and Tables	Probability
Usage	To describe a situation.	To explain the chances of occurrence of an event.
Function	It explains the data, which is already known, to summarize sample.	It attempts to reach the conclusion to learn about the population, that extends beyond the data available.

## 2. Differentiate between population and sample?

BASIS FOR COMPARISON	POPULATION	SAMPLE
Meaning	Population refers to the collection of all elements possessing common characteristics, that comprises universe.	Sample means a subgroup of the members of population chosen for participation in the study.
Includes	Each and every unit of the group.	Only a handful of units of population.
Characteristic	Parameter	Statistic
Data collection	Complete enumeration or census	Sample survey or sampling
Focus on	Identifying the characteristics.	Making inferences about population.

## 3. What is hypothesis? Differentiate between null and alternative hypothesis?

- Hypothesis is the assumption about the population parameter. The assumption made by the research may or may not be true.
- It is the fixed relationship between two or more variables which direct the research activity to test it.
- It should be testable either by experiment or observation.

Differentiate between null hypothesis and alternative hypothesis:

BASIS FOR COMPARISON	NULL HYPOTHESIS	ALTERNATIVE HYPOTHESIS
Meaning	A null hypothesis is a statement, in which there is no relationship between two variables.	An alternative hypothesis is a statement in which there is some statistical significance between two measured phenomenon.
Represents	No observed effect	Some observed effect
What is it?	It is what the researcher tries to disprove.	It is what the researcher tries to prove.
Acceptance	No changes in opinions or actions	Changes in opinions or actions
Testing	Indirect and implicit	Direct and explicit
Observations	Result of chance	Result of real effect
Denoted by	H-zero	H-one
Mathematical formulation	Equal sign	Unequal sign

#### 4. What is the central limit theorem?

- The central limit theorem states that if you take sufficiently large samples from a population, the samples' means will be normally distributed, even if the population isn't normally distributed.
- The central limit theorem is useful when analyzing large data sets because it allows one to assume that the sampling distribution of the mean will be normally-distributed in most cases. This allows for easier statistical analysis and inference.
- The central limit theorem links the following two distributions: The distribution of the variable in the population. The sampling distribution of the mean.
- Most common measure of central limit theorem are mean, median, mode.

#### 5. Differentiate between type-1 and type-2 error?

##### Type -1 Error (Error of the first kind)

- It is also known as a false-positive.
- It occurs if the researcher rejects a correct null hypothesis in the population.
  - i.e., incorrect rejection of the null hypothesis.
- Measured by alpha (significance level).
- If the significance level is fixed at 5%,
  - It means there are about five chances of type – 1 error out of 100.
- Cause of Type – 1 Error
  - The significance level is decided before testing the hypothesis
  - Sample size is not considered
  - This may occur due to chance
- It can be reduced by decreasing the level of significance.

##### Type -2 Error (Error of the second kind)

- It is also known as a false negative.
- It occurs if a researcher fails to reject a null hypothesis that is actually a false hypothesis.
- Measured by beta (the power of test).
- The probability of committing a type -2 error is calculated by  $1 - \text{beta}$  (the power of test).
- Cause of Type – 2 Error:
  - A statistical test is not powerful enough.
  - It is caused by a smaller sample size.
    - It may hide the significance level of the items being tested.
- It can be reduced by increasing the level of significance.

## 6. What is linear regression?

Linear regression analysis is used to predict the value of a variable based on the value of another variable. The variable you want to predict is called the dependent variable. The variable you are using to predict the other variable's value is called the independent variable.

## 7. What are the assumptions required for linear regression?

Five main assumptions underlying multiple regression models must be satisfied:

- (1) linearity
- (2) homoskedasticity
- (3) independence of errors
- (4) normality
- (5) independence of independent variables. Diagnostic plots can help detect whether these assumptions are satisfied.

