1) **Statement of Central limit Theorem:**
The central limit theorem states that if we have a population with mean μ and standard deviation σ and take sufficiently large random samples from the population with replacement, then the distribution of the sample mean is asymptotically normal.
We can calculate the mean of the sample means for the random samples we choose from the population:
$\mu_{\bar{X}}=\mu$
As well as the standard deviation of sample means:
$\sigma_{\bar{X}}=\sigma n$
According to the central limit theorem, the form of the sampling distribution will approach normalcy as the sample size is sufficiently large (usually n>30). regardless of the population distribution.
Importance of Central Limit Theorem:
This is useful since the researcher never knows which mean in the sampling distribution corresponds to the population mean, but by taking numerous random samples from a population, the sample means will cluster together, allowing the researcher to obtain a very accurate estimate of the population mean.

2) In Statistics, the **sampling method** or **sampling technique** is the process of studying the population by gathering information and analyzing that data. It is the basis of the data where the sample space is enormous.

There are several different sampling techniques available, and they can be subdivided into two groups. All these methods of sampling may involve specifically targeting hard or approach to reach groups.

Types of Sampling Method

In Statistics, there are different sampling techniques available to get relevant results from the population. The two different types of sampling methods are::

- Probability Sampling
- Non-probability Sampling

3) Type I Error

A type I error appears when the null hypothesis (H_0) of an experiment is true, but still, it is rejected. It is stating something which is not present or a false hit. A type I error is often called a false positive (an event that shows that a given condition is present when it is absent). In words of community tales, a person may see the bear when there is none (raising a false alarm) where the null hypothesis (H_0) contains the statement: “There is no bear”.

The type I error significance level or rate level is the probability of refusing the null hypothesis given that it is true. It is represented by Greek letter α (alpha) and is also known as alpha level. Usually, the significance level or the

probability of type i error is set to 0.05 (5%), assuming that it is satisfactory to have a 5% probability of inaccurately rejecting the null hypothesis.

Type II Error

A type II error appears when the null hypothesis is false but mistakenly fails to be refused. It is losing to state what is present and a miss. A type II error is also known as false negative (where a real hit was rejected by the test and is observed as a miss), in an experiment checking for a condition with a final outcome of true or false.

A type II error is assigned when a true alternative hypothesis is not acknowledged.

4) The Normal Distribution is defined by the probability density function for a continuous random variable in a system. Let us say, f(x) is the probability density function and X is the random variable. Hence, it defines a function which is integrated between the range or interval (x to x + dx), giving the probability of random variable X, by considering the values between x and x+dx.

$f(x) \geq 0 \; \forall \; x \in (-\infty, +\infty)$

And $\int_{-\infty}^{+\infty} f(x) = 1$

Normal Distribution Formula

The probability density function of normal or gaussian distribution is given by;

$$f(x, \mu, \sigma) = \frac{1}{\sigma\sqrt{2\pi}} e^{\frac{-(x-\mu)^2}{2\sigma^2}}$$

.5)

Correlation & Covariance Comparative

Basis		Covariance		Correlation
Meaning		Covariance indicates the extent of the variable being dependent on each other. Higher value denotes higher dependency.		Correlation signifies the strength of association between the variables when the other things are constant.

Relationship		Correlation can be gathered from covariance.		Correlation gives the value of covariance on a standard scale.
Values		Lie between -∞ and +∞		Correlation has limited values in the range of -1 and +1.
Scalability		Affects covariance		Correlation isn't affected by a change in scale.
Units		Covariance will have a definite unit as it is concluded from the multiplication of numbers and their units.		Correlation is a number without units but includes decimal values.

6) Univarate Analysis

Univariate analysis is the simplest form of data analysis where the data being analyzed contains only one variable. Since it's a single variable it doesn't deal with causes or relationships. The main purpose of univariate analysis is to describe the data and find patterns that exist within it.
You can think of the variable as a category that your data falls into. One example of a variable in univariate analysis might be "age". Another might be "height". Univariate analysis would not look at these two variables at the same time, nor would it look at the relationship between them.
Some ways you can describe patterns found in univariate data include looking at mean, mode, median, range, variance, maximum, minimum, quartiles, and standard deviation. Additionally, some ways you may display univariate data include frequency distribution tables, bar charts, histograms, frequency polygons, and pie charts.

Bivarate Analysis

Bivariate analysis is used to find out if there is a relationship between two different variables. Something as simple as creating a scatterplot by plotting one variable against another on a Cartesian plane (think X and Y axis) can sometimes give you a picture of what the data is trying to tell you. If the data seems to fit a line or curve then there is a relationship or correlation between the two variables. For example, one might choose to plot caloric intake versus weight.

Multivariate Analysis

Multivariate analysis is the analysis of three or more variables. There are many ways to perform multivariate analysis depending on your goals. Some of these methods include:

- Additive Tree
- Canonical Correlation Analysis
- Cluster Analysis
- Correspondence Analysis / Multiple Correspondence Analysis
- Factor Analysis
- Generalized Procrustean Analysis
- MANOVA
- Multidimensional Scaling
- Multiple Regression Analysis
- Partial Least Square Regression
- Principal Component Analysis / Regression / PARAFAC
- Redundancy Analysis.

1. 7) The sensitivity is calculated by dividing the percentage change in output by the percentage change in input.

This process of testing sensitivity for another input (say cash flows growth rate) while keeping the rest of inputs constant is repeated until the sensitivity figure for each of the inputs is obtained. The conclusion would be that the higher the sensitivity

Methods of Sensitivity Analysis

There are different methods to carry out the sensitivity analysis:

- Modeling and simulation techniques
- Scenario management tools through Microsoft excel

There are mainly two approaches to analyzing sensitivity:

- Local Sensitivity Analysis
- Global Sensitivity Analysis

8) *Hypothesis testing ascertains whether a particular assumption is true for the whole population. It is a statistical tool. It determines the validity of inference by evaluating sample data from the overall population.*
The concept of hypothesis works on the probability of an event's occurrence. It confirms whether the primary hypothesis results are correct or not. It is widely applied in research—biology, criminal trials, marketing, and manufacturing.

- Hypothesis testing is a statistical interpretation that examines a sample to determine whether the results stand true for the population.

- The test allows two explanations for the data—the null hypothesis or the alternative hypothesis. If the sample mean matches the population mean, the null hypothesis is proven true.

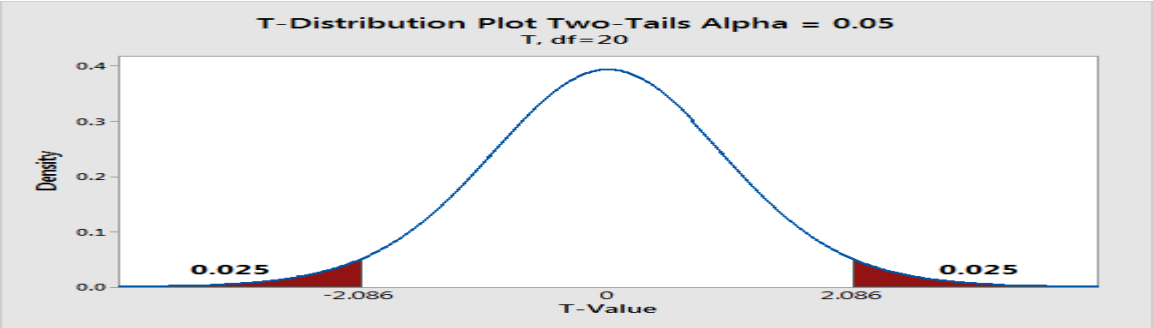
- Alternatively, if the sample mean is not equal to the population mean, the alternate hypothesis is accepted.

- This method requires superior analytical abilities and, therefore, is inaccessible for most. Also, this method heavily relies on probability.

Tails in a Hypothesis Test

These test statistics follow a sampling distribution. Probability distribution plots display the probabilities of obtaining test statistic values when the null hypothesis is correct. On a probability distribution plot, the portion of the shaded area under the curve represents the probability that a value will fall within that range.

The graph below displays a sampling distribution for t-values. The two shaded regions cover the two-tails of the distribution.



Two-Tailed Hypothesis Tests

Two-tailed hypothesis tests are also known as nondirectional and two-sided tests because you can test for effects in both directions. When you perform a two-tailed test, you split the significance level percentage between both tails of the distribution. In the example below, I use an alpha of 5% and the distribution has two shaded regions of 2.5% (2 * 2.5% = 5%).

9)