1. Statistical Distributions:

Discrete probability distributions:

In discrete case we have fine values, and we are getting probability at a particular point.
List of distributions: Discrete uniform distribution, Binomial, Poisson, Negative Binomial, Geometric, and so on.

Binomial Distribution: It is a generalized case of Bernoulli Distribution. In Bernoulli distribution random experiment whose outcomes are two types either success or failure, occurring with prob p and q respectively. In Binomial case we are repeating random experiments by n times, where n must be finite.

Continuous probability distributions:

List of continuous distributions: Normal, gamma, beta, exponential, laplace, weibull, logistics and so on.

Normal Distribution: It plays a very important role in statistics. In normal we have two parameters mean mu and variance sigma. If the value of n is too large in binomial distributions and values of p and q is very small then Binomial tends to normal.
Key points:
1. Probability curve is bell shaped.
2. mean , median and mode are same.
3. For large samples most of the distributions tends to normal
4. Even variable is not normal we can tend it to normal by simply transformation.
5. Entire theory of testing of hypotheses, linear regression, anova, based on assumptions of normal distribution.

2. Hypotheses of Testing:

Hypothesis testing is a crucial procedure to perform when you want to make inferences about a population using a random sample. These inferences include estimating population properties such as the mean, differences between means, proportions, and the relationships between variables.
In most cases, it is simply impossible to observe the entire population to understand its properties, so in such situations we go with samples. samples are much more practical and less expensive to work with.

Hypothesis testing is a statistical analysis that uses sample data to assess two mutually exclusive theories about the properties of a population. Statisticians call these theories the null hypothesis and the alternative hypothesis. A hypothesis test assesses your sample statistic and factors in an estimate of the sample error to determine which hypothesis the data support.

A parameter is a value that describes a characteristic of an entire population, such as the population mean.

Null Hypothesis
The null hypothesis is one of two mutually exclusive theories about the properties of the population in hypothesis testing. Typically, the null hypothesis states that there is no effect (i.e., the effect size equals zero). The null is often signified by H0.

Alternative Hypothesis
Typically, the alternative hypothesis states that a population parameter does not equal the null hypothesis value. H1

For example, in a 2-sample t-test, the alternative often states that the difference between the two means does not equal zero.

P-values
P-values are the probability that you would obtain the effect observed in your sample, or larger, if the null hypothesis is correct. In simpler terms, p-values tell you how strongly your sample data contradict the null.

Significance Level (Alpha)
The significance level, also known as alpha or α. It specifies how strongly the sample evidence must contradict the null hypothesis before you can reject the null for the entire population. **This standard is defined by the probability of rejecting a null hypothesis that is true.**

If the p-value is less than your significance level, you can reject the null and conclude that the effect is statistically significant.

Types of Errors in Hypothesis Testing
There are two types of errors related to drawing an incorrect conclusion.

- False positives: You reject a null that is true. Statisticians call this a Type I error. The Type I error rate equals your significance level or alpha (α).
- False negatives: You fail to reject a null that is false. Statisticians call this a Type II error. Generally, you do not know the Type II error rate. However, it is a larger risk when you have a small sample size, noisy data, or a small effect size. The type II error rate is also known as beta (β).

Tests:
1. One sample t test:
Test whether the mean of a normally distributed population is different from the specified values.
Null hypothesis: population mean is equal to specified value(mu0).

2. Two sample t test:

Test whether the means of two populations are same or differ.

ANOVA:
ANOVA is used when we want to check or compare means of more than two groups.