1. What is central limit theorem and why is it important?

Ans: Central limit theorem says that at least if we take 30 samples in the given data,It follows the normal distribution. which is also called as Bell shaped Curve.

It is really important as it allows to find the accurate probability of the data ,as we have taken adequate samples.

2. What is sampling? How many sampling methods do you know?

Ans: Sampling ,means we actually collect the group of data for research. For example we need to give opinion about the student in the university, we will take at least a sample of 1000 students.

We have 5 basic Sample methods:

- Simple Random.
- Convenience.
- Systematic.
- Cluster.
- Stratified.

3. What is the difference between type1 and typeII error?

Type 1 Error:

H0=Null Hypothesis.

HO is True and you reject it, it is called Type 1 Error.

Type 2 Error:

Ho is False and you failed to reject it ,it is called as Type 2 Error.

4.What do you understand by the term Normal distribution?

Ans: It is the distribution of the Dataset(Most Random Variables),follows a bell shaped curve Called as normal distribution.99.7 % of the dataset lies between 3 standard deviation of the mean.

5. What is correlation and covariance in statistics?

Ans Corelation: A measure which determines change in one variable due to change in one variable we have Positive corelation and negative corelation.

Positive corelation: if you study 1 hr /day your marks increases ,positive corelation.

Negative corelation: If you see more socialmedia/day your marks decreases ,negative corelation.

Covariance: change in variable reciprocated to equivalent change in another variable.

The value is from – 1 to +1.

If all the observation remans without any change then the covariance is zero.

6. Differentiate between univariate , Biavariate, and multivariate analysis

Univariate : Analysis on one variable.

Bivariate : Analysis on two variables.

Multivariate: Analysis based on more than two variables.

7. What do you understand by sensitivity and how would you calculate it?

Sensitivity:

How many Positive, correctly recalled by our model.

Sensitivity=TP/TP+FN.

8. What is hypothesis testing? What is H0 and H1? What is H0 and H1 for two-tail test?

Ans: Hypothesis Testing is a type of statistical analysis in which you put your assumptions about a population parameter to the test. It is used to estimate the relationship between 2 statistical variables.

The Null Hypothesis is the assumption that the event will not occur. A null hypothesis has no bearing on the study's outcome unless it is rejected.It is called as H0.

H1 is alternate hypothesis ,which is opposite to the Null hypothesis.

Two Tail test: In two tails, the test sample is checked to be greater or less than a range of values in a Two-Tailed test, implying that the critical distribution area is two-sided.

9. What is quantitative data and qualitative data?

Ans: Quantitative data are measures of values or counts and are expressed as numbers. Quantitative data are data about numeric variables (e.g. how many; how much; or how often). Qualitative data are measures of 'types' and may be represented by a name, symbol, or a number code.

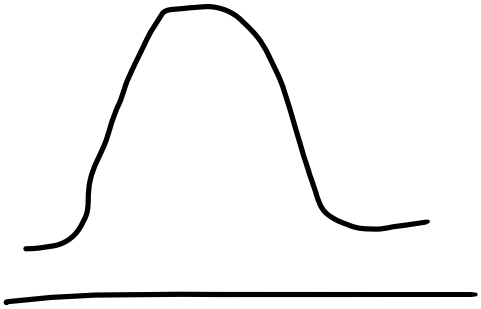
10. How to calculate range and interquartile range?

Ans: Range is calculated by substracting the lowest value from the Highest Value.

IQR is calculated by Q3-Q1,where q3 is 75% of the data present and q1 is 25% of the data present.

11. What do you understand by bell curve distribution ?

Ans: Bell Curve Distribution,is also called as normal distribution, if we take sample of the data and graph it follows a bell shape curve



The data is distributed in this shape.

12. Mention one method to find outliers:

Ans: By Using Boxplot, we can find outliers.

13. What is p-value in hypothesis testing?

Ans: P Value tells us the probability of finding the particular test of the observation to decide whether to reject the null hypothesis ho.

14. What is the Binomial Probability Formula?

The binomial distribution model allows us to compute the probability of observing a specified number of "successes" when the process is repeated a specific number of times (e.g., in a set of patients) and the outcome for a given patient is either a success or a failure.

15.Explain ANOVA and it's applications

Ans: An **ANOVA** test is a way to find out if survey or experiment results are significant. In other words, they help you to figure out if you need to reject the null hypothesis or accept the alternate hypothesis.

Basically, **you're testing groups to see if there's a difference between them.** Examples of when you might want to test different groups:

- A group of psychiatric patients are trying three different therapies: counseling, medication and biofeedback. You want to see if one therapy is better than the others.
- A manufacturer has two different processes to make light bulbs. They want to know if one process is better than the other.
- Students from different colleges take the same exam. You want to see if one college outperforms the other.

One-way or **two-way** refers to the number of independent variables (IVs) in your Analysis of Variance test.

- One-way has one independent variable (with 2 levels). For example: *brand of cereal*,
- Two-way has two independent variables (it can have multiple levels). For example: *brand of cereal, calories*.