Statistics Advance-2 Theory

1. What is hypothesis testing in statistics?

Hypothesis testing is a statistical method used to make decisions or inferences about population parameters based on sample data. It helps determine whether an observed effect is statistically significant.

2. What is the null hypothesis, and how does it differ from the alternative hypothesis?

- **Null Hypothesis (H_0):** Assumes no effect or difference. It's the default or status quo.

- **Alternative Hypothesis (H_1):** Represents the claim you're testing, indicating a significant effect or difference.

3. What is the significance level in hypothesis testing, and why is it important?

The **significance level (α)** is the probability of rejecting the null hypothesis when it is actually true. Common values: 0.05, 0.01. It sets the threshold for statistical significance.

4. What does a P-value represent in hypothesis testing?

The **P-value** is the probability of obtaining a test statistic as extreme as, or more extreme than, the observed result, assuming the null hypothesis is true.

5. How do you interpret the P-value in hypothesis testing?

- If $P \leq \alpha \rightarrow$ Reject H_0 (significant result).

- If $P > \alpha \rightarrow$ Fail to reject H_0 (not significant).

6. What are Type 1 and Type 2 errors in hypothesis testing?

- **Type I Error (False Positive):** Rejecting H_0 when it's true (controlled by α).

- **Type II Error (False Negative):** Failing to reject H_0 when it's false.

7. What is the difference between a one-tailed and a two-tailed test in hypothesis testing?

- **One-tailed test:** Tests for effect in one direction (e.g., greater than).

Statistics Advance-2 Theory

- **Two-tailed test:** Tests for effect in both directions (e.g., not equal to).

8. What is the Z-test, and when is it used in hypothesis testing?

The **Z-test** is used to test hypotheses about population means or proportions when the population standard deviation is known and sample size is large (n ≥ 30).

9. How do you calculate the Z-score, and what does it represent in hypothesis testing?

$Z = (\bar{X} - \mu) / (\sigma / \sqrt{n})$
 It tells how many standard deviations a sample mean is from the population mean under H_0.

10. What is the T-distribution, and when should it be used instead of the normal distribution?

The **T-distribution** is used when the population standard deviation is unknown and sample size is small (n < 30). It is wider and more spread out than the normal distribution.

11. What is the difference between a Z-test and a T-test?

- **Z-test:** Known population standard deviation, large sample.

- **T-test:** Unknown population standard deviation, small sample.

12. What is the T-test, and how is it used in hypothesis testing?

The **T-test** assesses whether the means of two groups are statistically different. Types include one-sample, two-sample (independent), and paired T-tests.

13. What is the relationship between Z-test and T-test in hypothesis testing?

Both test means, but differ in assumptions about sample size and population variance. T-test approximates Z-test as sample size increases.

14. What is a confidence interval, and how is it used to interpret statistical results?

A **confidence interval (CI)** gives a range of plausible values for a population parameter. A 95% CI means there's 95% confidence the true parameter lies within that range.

Statistics Advance-2 Theory

15. What is the margin of error, and how does it affect the confidence interval?

Margin of error is the range added/subtracted from the point estimate to create the CI. A larger margin increases the interval's width, indicating more uncertainty.

16. How is Bayes' Theorem used in statistics, and what is its significance?

Bayes' Theorem updates the probability of a hypothesis based on new evidence. It's crucial in decision-making, machine learning, and diagnostic testing.

17. What is the Chi-square distribution, and when is it used?

A **Chi-square distribution** is used in tests of categorical data, like independence tests or goodness-of-fit. It's skewed and only positive.

18. What is the Chi-square goodness of fit test, and how is it applied?

It checks if a sample matches an expected distribution. You compare observed vs. expected frequencies and use the Chi-square statistic to test fit.

19. What is the F-distribution, and when is it used in hypothesis testing?

The **F-distribution** is used to compare two variances. It's the basis of ANOVA and the F-test, and it's asymmetric and non-negative.

20. What is an ANOVA test, and what are its assumptions?

ANOVA (Analysis of Variance) tests if there are significant differences between three or more group means.
 Assumptions:

- Normality

- Homogeneity of variance

- Independence of observations

21. What are the different types of ANOVA tests?

- **One-way ANOVA:** One independent variable

<u>Statistics Advance-2 Theory</u>

- **Two-way ANOVA:** Two independent variables

- **Repeated Measures ANOVA:** Same subjects measured multiple times

22. What is the F-test, and how does it relate to hypothesis testing?

The **F-test** compares two variances to determine if they are significantly different. It's often used in ANOVA to determine if at least one group mean is different.