STATISTICS WORKSHEET-10

Q1 to Q12 have only one correct answer. Choose the correct option to answer your question.

1. Rejection of the null hypothesis is a conclusive proof that the alternative hypothesis is
a. True
b. False
c. Neither

2. Parametric test, unlike the non-parametric tests, make certain assumptions about
a. The population size
b. The underlying distribution
c. The sample size

3. The level of significance can be viewed as the amount of risk that an analyst will accept when making a decision
a. True
b. False

4. By taking a level of significance of 5% it is the same as saying
a. We are 5% confident the results have not occurred by chance
b. We are 95% confident that the results have not occurred by chance
c. We are 95% confident that the results have occurred by chance

5. One or two tail test will determine
a. If the two extreme values (min or max) of the sample need to be rejected
b. If the hypothesis has one or possible two conclusions
c. If the region of rejection is located in one or two tails of the distribution

6. Two types of errors associated with hypothesis testing are Type I and Type II. Type II error is committed when
a. We reject the null hypothesis whilst the alternative hypothesis is true
b. We reject a null hypothesis when it is true
c. We accept a null hypothesis when it is not true

7. A randomly selected sample of 1,000 college students was asked whether they had ever used the drug Ecstasy. Sixteen percent (16% or 0.16) of the 1,000 students surveyed said they had. Which one of the following statements about the number 0.16 is correct?
a. It is a sample proportion.
b. It is a population proportion.
c. It is a margin of error.
d. It is a randomly chosen number.

8. In a random sample of 1000 students, pˆ = 0.80 (or 80%) were in favour of longer hours at the school library. The standard error of pˆ (the sample proportion) is
a. .013
b. .160
c. .640
d. .800

9. For a random sample of 9 women, the average resting pulse rate is x = 76 beats per minute, and the sample standard deviation is s = 5. The standard error of the sample mean is
a. 0.557
b. 0.745
c. 1.667
d. 2.778

10. Assume the cholesterol levels in a certain population have mean μ= 200 and standard deviation σ = 24. The cholesterol levels for a random sample of n = 9 individuals are measured and the sample mean x is determined. What is the z-score for a sample mean x = 180?
a. –3.75
c. –2.50
c. −0.83
d. 2.50

11. In a past General Social Survey, a random sample of men and women answered the question "Are you a member of any sports clubs?" Based on the sample data, 95% confidence intervals for the population proportion who would answer "yes" are .13 to .19 for women and .247 to .33 for men. Based on these results, you can reasonably conclude that
a. At least 25% of American men and American women belong to sports clubs.
b. At least 16% of American women belong to sports clubs.
c. There is a difference between the proportions of American men and American women who belong to sports clubs.
d. There is no conclusive evidence of a gender difference in the proportion belonging to sports clubs.

12. Suppose a 95% confidence interval for the proportion of Americans who exercise regularly is 0.29 to 0.37. Which one of the following statements is FALSE?
a. It is reasonable to say that more than 25% of Americans exercise regularly.
b. It is reasonable to say that more than 40% of Americans exercise regularly.
c. The hypothesis that 33% of Americans exercise regularly cannot be rejected.
d. It is reasonable to say that fewer than 40% of Americans exercise regularly.

Q13 to Q15 are subjective answers type questions. Answers them in their own words briefly.
13. How do you find the test statistic for two samples?
To find the test statistic for two samples:
t = (x̄1 - x̄2) / (s√(1/n1 + 1/n2))

where:
x̄1 and x̄2 are the sample means of the two samples
s is the pooled standard deviation, calculated as:
s = √((s1^2(n1-1) + s2^2(n2-1)) / (n1 + n2 - 2))

s1 and s2 are the standard deviations of the two samples
n1 and n2 are the sample sizes of the two samples
The test statistic will follow a t-distribution with (n1 + n2 - 2) degrees of freedom. Once you have the test statistic, you can compare it to the critical value from the t-distribution or use it to calculate the p-value to make a decision about the null hypothesis.

14. How do you find the sample mean difference?

To find the sample mean difference between two samples, you subtract the mean of one sample from the mean of the other sample. The formula for calculating the sample mean difference is:

sample mean difference = $\bar{x}1 - \bar{x}2$

Where:

$\bar{x}1$ is the mean of the first sample

$\bar{x}2$ is the mean of the second sample

For example, suppose you have two samples of exam scores, with a mean score of 75 in Sample 1 and a mean score of 80 in Sample 2. The sample mean difference would be:

sample mean difference = 75 - 80 = -5

This means that the mean score in Sample 1 is 5 points lower than the mean score in Sample 2.

15. What is a two sample t test example?

A two-sample t-test is a statistical test used to compare the means of two independent groups. An example of a two-sample t-test is:

Suppose we want to compare the mean weight of apples from two different farms, Farm A and Farm B. We randomly select 10 apples from Farm A and 10 apples from Farm B and measure their weights in grams. The weights for Farm A are: 150, 152, 155, 148, 147, 151, 153, 149, 155, 147. The weights for Farm B are: 158, 157, 160, 155, 162, 156, 159, 154, 158, 161.

To determine if there is a significant difference in the mean weight of apples from the two farms, we can perform a two-sample t-test. We can set up the null hypothesis as:

$H0: \mu A = \mu B$

where μA is the population mean weight of apples from Farm A, and μB is the population mean weight of apples from Farm B.

The alternative hypothesis can be either:

$Ha: \mu A < \mu B$ (left-tailed test)

$Ha: \mu A \neq \mu B$ (two-tailed test)

$Ha: \mu A > \mu B$ (right-tailed test)

We can then calculate the sample mean weight difference (also known as the point estimate) as:

$\bar{x}A - \bar{x}B = (150 + 152 + 155 + 148 + 147 + 151 + 153 + 149 + 155 + 147)/10 - (158 + 157 + 160 + 155 + 162 + 156 + 159 + 154 + 158 + 161)/10 = -5.2$

This means that, on average, the apples from Farm A are 5.2 grams lighter than the apples from Farm B.

We also need to calculate the pooled standard deviation (s) and the t-statistic using the formula:

s = sqrt[((n1-1)s1^2 + (n2-1)s2^2) / (n1+n2-2)]

t = $(\bar{x}A - \bar{x}B)$ / [s * sqrt(1/n1 + 1/n2)]

where n1 and n2 are the sample sizes, s1 and s2 are the sample standard deviations, and $\bar{x}A$ and $\bar{x}B$ are the sample means.

Using a significance level of 0.05 and degrees of freedom of (n1+n2-2) = 18, we can then determine the critical t-value from a t-distribution table or a calculator. If the calculated t-value is greater than the critical t-value, we can reject the null hypothesis and conclude that there is a significant difference in the mean weight of apples from the two farms.