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?

Answer:- The test statistic for comparing two samples depends on the specific hypothesis test being conducted. However, in general, the test statistic is calculated by taking the difference between the sample means (or proportions) and dividing by the standard error of the difference.

For example, in a two-sample t-test, the test statistic is calculated as:

$$t = (\bar{x}_1 - \bar{x}_2) / [s^2_{pooled} / (n_1 + n_2)]$$

where $\bar{x}_1$ and $\bar{x}_2$ are the sample means of the two groups, s^2_{pooled} is an estimate of the common variance of the two groups, and n_1 and n_2 are the sample sizes.

In a two-sample z-test for proportions, the test statistic is calculated as:

$$z = (\hat{p}_1 - \hat{p}_2) / \mathrm{sqrt}\{ \hat{p}(1 - \hat{p}) [(1/n_1) + (1/n_2)] \}$$

where $\hat{p}_1$ and $\hat{p}_2$ are the sample proportions of the two groups, $\hat{p}$ is the pooled proportion of the two groups, and n_1 and n_2 are the sample sizes.

These are just two examples of how to calculate the test statistic for two-sample hypothesis tests. The specific formula will depend on the type of test and the specific hypothesis being tested.

14. How do you find the sample mean difference?

Answer:- The sample mean difference (denoted as $\bar{d}$) is found by taking the difference between the sample means of two samples. Mathematically, it is expressed as:

$$\bar{d} = \bar{x}_1 - \bar{x}_2$$

where $\bar{x}_1$ is the sample mean of the first sample and $\bar{x}_2$ is the sample mean of the second sample.

15. What is a two sample t test example?

Answer:- A two-sample t-test is used to compare the means of two independent groups or samples. Here is an example:

Suppose we want to compare the average height of men and women. We collect two random samples, one consisting of 50 men and the other consisting of 50 women. We measure their heights and calculate the sample means and standard deviations as follows:

Sample mean height for men ($\bar{x}_1$): 180 cm

Sample standard deviation for men (s_1): 7 cm

Sample mean height for women ($\bar{x}_2$): 165 cm

Sample standard deviation for women (s_2): 6 cm

We can use a two-sample t-test to determine if the difference in the sample means is statistically significant. The null hypothesis is that the population means are equal, while the alternative hypothesis is that they are not equal. We can calculate the test statistic as follows:

$t = \frac{(\bar{x}_1 - \bar{x}_2) - (\mu_1 - \mu_2)}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}}$

where μ_1 and μ_2 are the population means (unknown), n_1 and n_2 are the sample sizes, and $\bar{x}_1 - \bar{x}_2$ is the difference in sample means.

If the calculated t-value exceeds the critical t-value at a certain level of significance and degrees of freedom, we reject the null hypothesis and conclude that there is a statistically significant difference between the population means.