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

Answer
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

Answer
b. The underlying distribution

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

Answer
a. True

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

Answer
b.We are 95% confident that the results have not 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

Answer
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

Answer
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.

Answer
a. It is a sample proportion.

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

Answer
a. .013

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
Answer
c.1.667

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
Answer
b. –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.

Answer
C. There is a difference between the proportions of American men and American women who belong 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.

Answer
b.It is reasonable to say that more 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 two-sample t-test is one of the most common statistical tests used. It is applied to compare whether the averages of two data sets are significantly different, or if their difference is due to random chance alone. It could be used to determine if a new teaching method has really helped teach a group of kids better, or if that group is just more intelligent. .

Here are some common methods for comparing two samples:

1. Two-sample t-test: If you want to test whether the means of two populations are equal, you can use a two-sample t-test. The test statistic for this test is: **$t = (\bar{x}_1 - \bar{x}_2) / (s_1^2/n_1 + s_2^2/n_2)$^(1/2)**

2. Wilcoxon signed-rank test: If you have paired data and want to compare the medians of the differences between the pairs, you can use the Wilcoxon signed-rank test. The test statistic for this test is based on the sum of the signed ranks of the differences.

3. Mann-Whitney U test: If you want to compare the medians of two populations, you can use the Mann-Whitney U test. The test statistic for this test is**: $U = n_1 n_2 + (n_1(n_1 + 1))/2 - R_1$**

4. Paired t-test: If you have paired or matched observations, such as beforeand-after measurements, you can use a paired t-test. The test statistic for this test is: **$t = (\bar{x}d - \mu d) / (sd / \sqrt{n})$** where $\bar{x}d$ is the sample mean of the differences, μd is the hypothesized mean difference, sd is the sample standard deviation of the differences, and n is the sample size.

14. How do you find the sample mean difference?

Answer

To find the sample mean difference, you need to have two sets of data that you want to compare. Here are the steps to find the sample mean difference:

1. Calculate the difference between the two means by subtracting X2 from X1. This gives you the sample mean difference
2. Take the first set of data and calculate the mean (average) of that set. Let's call this value X1.
3. Take the second set of data and calculate the mean of that set. Let's call this value X2.

two sets of data

* Set 1: 10, 20, 30, 40, 50

*Set 2: 15, 25, 35, 45, 55

To find the sample mean difference between these two sets of data, you would do the following:

1. Calculate the mean of set 1: X1 = (10 + 20 + 30 + 40 + 50) / 5 = 30
2. Calculate the mean of set 2: X2 = (15 + 25 + 35 + 45 + 55) / 5 = 35
3. Calculate the sample mean difference: Sample mean difference = X1 - X2 = 30 - 35 = -5

the sample mean difference between these two sets of data is -5

15. What is a two sample t test example?

Answer

A two-sample t-test is a statistical hypothesis test used to determine whether two independent groups of data have different means. Here's an example:

Suppose you want to compare the average test scores of two different classes of students (Class A and Class B) to determine whether one class performed significantly better than the other. You randomly select 20 students from each class and record their test scores.

Here are the scores:

Class A: 85, 78, 92, 90, 84, 88, 76, 91, 82, 87, 89, 93, 79, 81, 83, 90, 87, 92, 88, 86

Class B: 76, 82, 85, 80, 72, 77, 83, 81, 78, 73, 79, 84, 81, 75, 80, 82, 86, 77, 79, 84

To perform a two-sample t-test, you first calculate the means and standard deviations of the two groups:

Class A: mean = 86.6, standard deviation = 5.81

Class B: mean = 79.2, standard deviation = 4.14

You can see that Class A has a higher mean score than Class B. However, you need to determine whether this difference is statistically significant. You can do this by calculating the t-statistic:

t = (mean(A) - mean(B)) / (sqrt(s^2(A)/n(A) + s^2(B)/n(B)))

where mean(A) and mean(B) are the means of Class A and Class B, s(A) and s(B) are the standard deviations of Class A and Class B, and n(A) and n(B) are the sample sizes.

Plugging in the numbers, we get:

t = (86.6 - 79.2) / (sqrt((5.81^2/20) + (4.14^2/20))) = 3.13

Next, we compare the t-value to a critical value from a t-distribution with (n(A) + n(B) - 2) degrees of freedom at a chosen significance level (e.g., $\alpha = 0.05$). If the calculated t-value is greater than the critical t-value, we reject the null hypothesis that the means are equal and conclude that the two groups have significantly different mean. In this case, the critical t-value is 2.086 at $\alpha = 0.05$ with 38 degrees of freedom. Since our calculated t-value of 3.13 is greater than 2.086, we reject the null hypothesis and conclude that there is a significant difference in test scores between Class A and Class B