STATISTICS WORKSHEET

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

Ans: c

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

Ans: b

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

Ans: a

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

Ans: b

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

Ans: c

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

Ans: c

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.

Ans: a

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

Ans: a

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

Ans: c

10. Assume the cholesterol levels in a certain population have mean $\mu = 200$ and standard deviation $\sigma = 24$. 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
b. –2.50
c. −0.83
d. 2.50

Ans: b

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.

Ans: d

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.

Ans: b

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?
Ans: The test statistic of two samples is find by performing independent t test at significance level 0.05. It is calculated by taking the difference between two sample means and dividing by the standard deviation.

14. How do you find the sample mean difference?
Ans: To find the sample mean difference first we need to calculate the mean of samples then subtract one mean from another.

15. What is a two sample t test example?

Ans: Here is an example of two sample t test:

Lets say we need to compare the performance of two call centers in terms of average call lengths and find out if the there is a difference between average call lengths or they are same. For that first we need to consider null and alternate hypothesis.

Null hypothesis, H0: There is no difference between the average call length between two call centers.
Alternate hypothesis, Ha: There is a difference between the average call length.

We randomly select 20 calls from each call center and calculate the average call lengths.

First, we need to calculate the two sample means and standard deviations:
Call Center A: Sample mean, **m1** = 122 seconds, SD, **S1** = 15 seconds, **n1** = 20
Call Center B: Sample mean, **m2** = 135 seconds, SD, **S2** = 20 seconds, **n2** = 20

Now, we use a two-sample t-test to determine if the difference between two sample means is statistically significant. We will use a 95% confidence level and alpha = 0.05.

The two-sample t-statistic is calculated as the following assuming that the standard deviations of the population is not same and the population mean is same.

t = ((135 – 122) – 0)/SQRT((20*20/20) + ((15*15)/20))

t = 13/SQRT(20 + 11.25)

t = 13/SQRT(31.25)

t = 2.3256
The value of degrees of freedom can be calculated as the following:

Degree of freedom, df = n1 + n2 -2 = 20 + 20 – 2 = 38

The critical value of a two-tailed T-test with degrees of freedom as 38 and level of significance as 0.05 comes out to be **2.0244**. Since the current t-value of 2.3256 is greater than the critical value of 2.0244, one can reject the null hypothesis that there is no difference between the performance in terms of the call length time. Thus, based on the given evidence, the alternate hypothesis stands as true. i.e., there is a difference between the average call length.