**Statistics Assignment: Confidence Interval and Hypothesis Testing**

Part A: Confidence Intervals

**1. Confidence Interval Calculation:**

* Given a sample mean of 50, a standard deviation of 10, and a sample size of 100, calculate the 95% confidence interval for the population mean.

--------------------------------------------------------------------------------------------------------------------

**Answer:**

To calculate the 95% confidence interval for the population mean, we use the following

formula:

Confidence Interval = Sample Mean± (Z-value × Standard Deviation Size/ √Sample Size​)

The Z-value for a 95% confidence interval is approximately 1.96

Confidence Interval= 50 ± (1.96×10/√100​)

Confidence Interval= 50 ± (1.96×1)

Confidence Interval= 50 ± 1.96

So, the 95% confidence interval for the population mean is approximately 48.0448.04 to 51.9651.96.

* Explain in your own words what this confidence interval represents.

**Answer:**

Example we are trying to find out the average height of all the people in a city, but we cannot measure everyone. So, we can take a sample, measure their heights, and calculate an average.

We can say that the true average height of everyone in the city is somewhere between 48.04 inches and 51.96 inches. The 95% confidence part means that if we did this process a bunch of times, we'd expect to capture the true average in that range about 95 out of 100 times.

So, it's our best guess for the average height.

**2. Interpretation of Results:**

* A biologist measures the length of 50 fish and finds a mean length of 20 cm with a standard deviation of 4 cm. Calculate the 90% confidence interval for the mean length of the fish population.
* ---------------------------------------------------------------------------------------------------------------------

**Answer:**

* To calculate the 90% confidence interval for the mean length of the fish population, we can use a similar formula as before:
* Confidence Interval=Sample Mean±(Z value × Standard Deviation /√Sample Size)

The Z-value for a 90% confidence interval is approximately 1.645

Confidence Interval = 20 ± (1.645 × 4 / √50​)

Confidence Interval=20±(1.645×0.566)

Confidence Interval=20±0.93

the 90% confidence interval for the mean length of the fish population is approximately 19.0719.07 cm to 20.9320.93 cm.

* Discuss what this interval implies about the actual mean length of the fish population.

**Answer:**

We are 90% confident that the true mean length of the fish population falls within the range of 19.0719.07 cm to 20.9320.93 cm

The 90% confidence level means that if we were to take many samples and calculate the confidence interval for each, about 90% of those intervals would capture the true mean length of the fish population.

If we need a higher level of confidence, we can choose a higher confidence level (e.g., 95% or 99%), but this would result in a wider interval.

Part B: Hypothesis Testing

**3. Hypothesis Test Setup:**

* A company claims that its product increases productivity by more than 15%. To test this claim, a sample of 30 employees is tested with a mean increase in productivity of 18% and a standard deviation of 5%. Perform a hypothesis test at the 5% significance level to see if the company's claim can be supported.
* State the null and alternative hypothesis, the test statistic, and your conclusion.

**Answer:**

Null Hypothesis (H0): The product does not increase productivity by more than 15%. HO: µ ≤ 15

Alternative Hypothesis (H1): The product increases productivity by more than 15%. H1 : µ ≥ 15

The significance level α is 5% or 0.05.

we can use a t-test because we are dealing with a sample and we don't know the population standard deviation

t = xˉ - µo/s/√n

xˉ is the sample mean (18%)

µo is the hypothesized population mean under the null hypothesis (15%)

s is the sample standard deviation (5%)

n is the sample size (30)

t = 18-15/5/√30

=3/5√30

=3x√30/5 = 2.75

The degrees of freedom (df) for this test is n−1=30−1=29

At the 5% significance level with 29 degrees of freedom, the critical value is approximately 1.699..

2.75 > 1.699

the calculated t-test statistic is greater than the critical value, we reject the null hypothesis

Conclusion: the claim that the company's product increases productivity by more than 15%. The sample data suggests a statistically significant increase in productivity.

**4. Analysis of Hypothesis Test:**

* A study claims that the average time spent on social media by teenagers is 3 hours a day. A researcher believes this has decreased due to new trends and conducts a study with 40 teenagers, finding an average time spent of 2.8 hours with a standard deviation of 0.5 hours. Conduct a hypothesis test at the 1% significance level.
* Discuss your findings and what they imply about the researcher's belief.
* **Answer:**

Claimed average time spent on social media by teenagers: 3 hours/day

Sample size (n) = 40

Sample mean (x̄) = 2.8 hours/day

Sample standard deviation (σ) = 0.5 hours/day

Null Hypothesis (H0): The average time spent on social media by teenagers is still 3 hours/day.

Alternative Hypothesis (H1): The average time spent on social media by teenagers has decreased from 3 hours/day.

Significance Level: 1% (α = 0.01)

Test Statistic for One-sample z-test: Z=​​xˉ−μ/σ/√n

xˉ is the sample mean

μ is the claimed population mean

σ is the population standard deviation (given as 0.5 hours)

n is the sample size

Z = 2.8-3/0.5/√40

-0.2/0.0791

Z = −2.53

With a significance level of 1% (α = 0.01), the critical z-value for a two-tailed test is approximately ±2.576.

The calculated z-value (-2.53) is less than -2.576.

The results of the hypothesis test provide sufficient evidence to reject the null hypothesis. Therefore, there is statistical significance at the 1% level, indicating that the average time spent on social media by teenagers has decreased from the claimed average of 3 hours per day. The sample of 40 teenagers, with an average time of 2.8 hours and a standard deviation of 0.5 hours, supports the researcher's belief that new trends have led to a reduction in the time teenagers spend on social media.

Part C: Reflection

**5. Real-World Application:**

* Choose a real-world scenario where confidence intervals and hypothesis testing could be applied. Describe the scenario and how these statistical methods could help in making decisions or drawing conclusions in that context.

**Drug Efficacy in Treating a Medical Condition**

Description: A pharmaceutical company has developed a new drug to treat a specific medical condition, such as hypertension. Before the drug is released to the market, the company wants to assess its efficacy compared to the current standard treatment. They conduct a clinical trial involving a sample of patients diagnosed with the medical condition.

**Application of Statistical Methods:**

**Confidence Intervals (CI):**

The researchers can use confidence intervals to estimate the range in which the true effect of the new drug is likely to fall. For instance, they might calculate a 95% confidence interval for the mean reduction in blood pressure associated with the new drug.

A narrow confidence interval would suggest more precise estimates, providing confidence in the drug's effectiveness. A wider interval might indicate more uncertainty.

**Hypothesis Testing:**

The null hypothesis (H0) could be that the new drug has no significant effect compared to the standard treatment (e.g., the mean reduction in blood pressure is the same for both treatments).

The alternative hypothesis (H1) would be that the new drug is effective in reducing blood pressure.

Statistical tests could be conducted to determine whether there is enough evidence to reject the null hypothesis in favor of the alternative, indicating that the new drug is indeed effective.

**Decision-Making and Conclusions:**

**Confidence Intervals:**

If the confidence interval for the mean reduction in blood pressure with the new drug does not include zero, it suggests a statistically significant effect. This information aids in decision-making about the drug's potential efficacy.

**Hypothesis Testing:**

If the hypothesis test results in rejecting the null hypothesis, it implies that the new drug has a significant effect on reducing blood pressure. This information is crucial for regulatory approval and decision-making about whether to proceed with the drug's market release.

In summary, confidence intervals and hypothesis testing in this scenario provide valuable insights into the effectiveness of the new drug compared to the standard treatment, assisting researchers and decision-makers in the pharmaceutical industry in making informed choices about drug development and market introduction.