Assignment

Q1: What is the difference between a t-test and a z-test? Provide an example scenario where you would
use each type of test.

Ans:The main difference between a t-test and a z-test lies in the information available about the population. Both tests are used to make inferences about population parameters based on sample data, but they are applicable under different conditions:

Z-Test:

- A z-test is appropriate when the population standard deviation (
- �
- σ) is known.
- It is used when the sample size is large (typically
- �≥ 30
- $n \geq 30$).
- The z-test statistic is calculated as
- �‾−���
- n
-
- σ
-
  - x
  - ‾
  - $-\mu$
  -
- , where
- �‾
  - x
  - ‾
- is the sample mean,
- �
- μ is the population mean,
- �
- σ is the population standard deviation, and
- �
- n is the sample size.
- Example Scenario: Suppose you have a large sample of exam scores from a population with a known standard deviation, and you want to test if the average

exam score of a new group is significantly different from the known population mean.

T-Test:

- A t-test is used when the population standard deviation (
- �
- σ) is unknown and must be estimated from the sample data.
- It is suitable for small sample sizes (typically
- �<30
- $n<30$).
- The t-test statistic is calculated as
- �‾−���
- n
-
- s
-
    - x
    - ‾
    - $-\mu$
    -
- , where
- �‾
    - x
    - ‾
- is the sample mean,
- �
- μ is the population mean,
- �
- s is the sample standard deviation, and
- �
- n is the sample size.
- Example Scenario: If you want to compare the average height of two different populations and you have small samples from each population, a t-test would be appropriate because you may not know the population standard deviations.

In summary, use a z-test when the population standard deviation is known and the sample size is large, and use a t-test when the population standard deviation is unknown or when dealing with small sample sizes. The choice between the two depends on the characteristics of the data and the assumptions you can make about the population.

Q2: Differentiate between one-tailed and two-tailed tests.
Ans:The distinction between one-tailed (one-sided) and two-tailed (two-sided) tests refers to the directionality of the hypothesis being tested and the area under the probability distribution curve where critical regions are located. Here's a breakdown of the differences:

One-Tailed Test:

- In a one-tailed test, the critical region is on only one side of the distribution curve.
- The hypothesis specifies a direction for the effect (e.g., greater than, less than), and the test is designed to determine if the sample statistic is significantly different in that specified direction.
- The null hypothesis (
- �0
- H
- 0
-
-) typically includes an equality sign (e.g.,
- $\geq$
- $\geq$,
- $\leq$
- $\leq$), and the alternative hypothesis (
- �1
- H
- 1
-
- or
- ��
- H
- a
-
-) specifies a directional inequality.
- There are two types of one-tailed tests: right-tailed (upper-tailed) and left-tailed (lower-tailed).
    - Right-tailed test: The critical region is in the right tail of the distribution. It tests if the sample statistic is significantly greater than a specified value.

    - Left-tailed test: The critical region is in the left tail of the distribution. It tests if the sample statistic is significantly less than a specified value.

Example:

- Null Hypothesis (
- �0
- H
- 0
-
-):
- �≤10
- $\mu \leq 10$ (population mean is less than or equal to 10)
- Alternative Hypothesis (
- �1
- H
- 1
-
-):
- �>10
- $\mu > 10$ (population mean is greater than 10)
- Critical Region: Right-tailed

Two-Tailed Test:

- In a two-tailed test, the critical region is split between both sides of the distribution curve.
- The hypothesis is open to the possibility of an effect in either direction (greater than or less than), and the test is designed to determine if the sample statistic is significantly different from the null hypothesis in either direction.
- The null hypothesis (
- �0
- H
- 0
-
-) typically includes an equality sign, and the alternative hypothesis (
- �1
- H

- 1
-
- or
- ��
- H
- a
-
-) specifies a non-directional inequality.
- The critical region is divided into two equal parts, one in each tail of the distribution.

Example:

- Null Hypothesis (
- �0
- H
- 0
-
-):
- �=50
- $\mu=50$ (population mean is equal to 50)
- Alternative Hypothesis (
- �1
- H
- 1
-
-):
- �≠50
- μ
- =50 (population mean is not equal to 50)
- Critical Region: Two-tailed

In summary, the choice between a one-tailed or two-tailed test depends on the specific research question and the directional hypothesis being tested. A one-tailed test is more sensitive to detecting effects in a specified direction, while a two-tailed test is more conservative and considers effects in either direction.

Q3: Explain the concept of Type 1 and Type 2 errors in hypothesis testing. Provide an example scenario for
each type of error.

Ans:In hypothesis testing, errors can occur when making decisions about the null hypothesis (
�0

H

0

) and the alternative hypothesis (

�1

H

1

or

��

H

a

). There are two types of errors: Type I error and Type II error.

Type I Error:

- Definition: Occurs when the null hypothesis is incorrectly rejected when it is actually true. It represents a false positive or a "alpha" error.
- Symbol: Denoted by
- �
- α, the significance level.
- Example Scenario:
    - Suppose a medical researcher is testing a new drug for effectiveness. The null hypothesis (
    - �0
    - H

- 0
-
-) might state that the drug has no effect, but in reality, it is true (the drug is ineffective). If, based on the sample data, the researcher incorrectly rejects the null hypothesis and concludes that the drug is effective, it would be a Type I error.

Type II Error:

- Definition: Occurs when the null hypothesis is incorrectly not rejected when it is actually false. It represents a false negative or a "beta" error.
- Symbol: Denoted by
- �
- β.
- Example Scenario:
  - Using the same example, suppose the null hypothesis (
  - �0
  - H
  - 0
  -
  -) states that the drug has no effect, but in reality, it is false (the drug is effective). If, based on the sample data, the researcher fails to reject the null hypothesis and concludes that the drug is ineffective, it would be a Type II error.

In summary:

- Type I Error (False Positive): Incorrectly rejecting a true null hypothesis.
  - Probability of Type I error:
  - �
  - α.
  - Controlled by the chosen significance level (
  - �
  - α).
  - Often set at 0.05 or 0.01.
- Type II Error (False Negative): Incorrectly failing to reject a false null hypothesis.
  - Probability of Type II error:
  - �
  - β.
  - Depends on the effect size, sample size, and variability in the data.
  - Power of a test (

- 1−�
- $1-\beta$) is the probability of correctly rejecting a false null hypothesis.

The balance between Type I and Type II errors is crucial in hypothesis testing. Adjusting the significance level (

�

α) affects the likelihood of Type I errors, but it also affects the power of the test and, consequently, the likelihood of Type II errors. Researchers often need to carefully choose the appropriate level of significance and sample size to manage the trade-off between these two types of errors based on the specific context of their study.

Q4: Explain Bayes's theorem with an example.
Ans:Bayes's theorem is a mathematical formula that describes the probability of an event, based on prior knowledge of conditions that might be related to the event. It's named after the Reverend Thomas Bayes, who introduced the theorem. The formula is often written as:

�(�|�)=�(�|�)⋅�(�)�(�)

$P(A|B)=$

$P(B)$

$P(B|A)\cdot P(A)$

Where:

- �(�|�)
- $P(A|B)$ is the probability of event A occurring given that event B has occurred (this is the posterior probability).
- �(�|�)
- $P(B|A)$ is the probability of event B occurring given that event A has occurred.
- �(�)
- $P(A)$ is the prior probability of event A.
- �(�)

- $P(B)$ is the prior probability of event B.

Bayes's theorem is particularly useful for updating probabilities based on new evidence. Let's illustrate Bayes's theorem with a classic example known as the "diagnostic test" scenario:

Example: Diagnostic Test for a Disease

Suppose there is a diagnostic test for a rare disease, and we want to determine the probability that an individual has the disease given a positive test result.

- �
- A: The individual has the disease.
- �
- B: The test result is positive.

Let's define the probabilities:

- �(�)
- $P(A)$: The prior probability of having the disease (without considering the test result).
- �(�|�)
- $P(B|A)$: The probability of testing positive given that the individual has the disease (sensitivity of the test).
- �(�)
- $P(B)$: The probability of testing positive (irrespective of whether the individual has the disease or not).
- �(�|�)
- $P(A|B)$: The probability of having the disease given a positive test result (posterior probability).

The formula can be applied as follows:

�(�|�)=�(�|�)·�(�)�(�)

$P(A|B)=$

$$P(B)$$

$$P(B \mid A) \cdot P(A)$$

- �(�|�)
- $P(A \mid B)$: The probability of having the disease given a positive test result.
- �(�|�)
- $P(B \mid A)$: The probability of testing positive given that the individual has the disease.
- �(�)
- $P(A)$: The prior probability of having the disease.
- �(�)
- $P(B)$: The probability of testing positive.

The example demonstrates how Bayes's theorem allows us to update our beliefs about the probability of having a disease based on the results of a diagnostic test, taking into account both the sensitivity of the test and the prior probability of the disease.

Q5: What is a confidence interval? How to calculate the confidence interval, explain with an example.
Ans:A confidence interval is a statistical range that is used to estimate the range in which a population parameter (such as the mean or proportion) is likely to lie. It provides a measure of the precision or uncertainty associated with a sample estimate. The interval is expressed as a range of values, and we say we are, for example, "95% confident" that the true parameter falls within that range.

The general formula for a confidence interval for the population mean (

�

μ) is given by:

Confidence Interval=(�¯−���,�¯+���)

Confidence Interval=(

$$x$$

‾

$-Z$

n

s

,

x

‾

$+Z$

n

s

)

Where:

- �‾
- x
- ‾
- is the sample mean.
- �
- s is the sample standard deviation.
- �
- n is the sample size.
- �

- Z is the Z-score corresponding to the desired confidence level.

Let's go through an example:

Example: Confidence Interval for Population Mean

Suppose we want to estimate the average height (in inches) of a population. We take a random sample of 30 individuals and find the sample mean (

�ˉ

x

ˉ

) to be 65 inches with a sample standard deviation (

�

s) of 3 inches.

Assuming a normal distribution and aiming for a 95% confidence interval, the Z-score for a 95% confidence level is approximately 1.96.

Confidence Interval=(65−1.96330,65+1.96330)

Confidence Interval=(65−1.96

30

3

,65+1.96

30

)

Now, let's calculate this:

python

Copy code

```
import

65
3
30
0.95

1.96

print f"95% Confidence Interval for Population Mean: ({lower_bound:.2f},
{upper_bound:.2f})"
```

In this example, the 95% confidence interval for the population mean height is approximately (64.25, 65.75). This means that we are 95% confident that the true average height of the population falls within this interval.

Q6. Use Bayes' Theorem to calculate the probability of an event occurring given prior knowledge of the event's probability and new evidence. Provide a sample problem and solution.

Ans:Let's consider a classic example involving a medical scenario: the probability of having a rare disease given a positive test result.

Example: Diagnostic Test for a Disease

Suppose there is a diagnostic test for a rare disease, and we want to determine the probability that an individual has the disease given a positive test result.

Let's define the events:

- �
- A: The individual has the disease.
- �
- B: The test result is positive.

We know the following probabilities:

- �(�)
- $P(A)$: The prior probability of having the disease (without considering the test result).
- �(�∣�)
- $P(B \mid A)$: The probability of testing positive given that the individual has the disease (sensitivity of the test).
- �(¬�)
- $P(\neg A)$: The complement of
- �(�)
- $P(A)$, which is the probability of not having the disease.
- �(�∣¬�)

- $P(B|\neg A)$: The probability of testing positive given that the individual does not have the disease (false positive rate).
- �(�)
- $P(B)$: The probability of testing positive.

Now, we can use Bayes's theorem to calculate the probability of having the disease given a positive test result:

�(�|�)=�(�|�)⋅�(�)�(�)

$P(A|B)=$

$$P(B)$$

$$P(B|A)\cdot P(A)$$

Let's assume the following probabilities:

- �(�)
- $P(A)$: 0.01 (1% of the population has the disease).
- �(�|�)
- $P(B|A)$: 0.95 (the test correctly identifies 95% of those with the disease).
- �(¬�)
- $P(\neg A)$: 0.99 (99% of the population does not have the disease).
- �(�|¬�)
- $P(B|\neg A)$: 0.05 (5% false positive rate).
- �(�)
- $P(B)$: Calculated using the law of total probability.

python

Copy code

0.01

0.95

0.99

0.05

```
print f"The probability of having the disease given a positive test result is:
{P_A_given_B:.4f}"
```

In this example, the calculated probability

�(�|�)

$P(A \mid B)$ represents the updated belief in the probability of having the disease given the new evidence of a positive test result. The actual values used in this example are for illustrative purposes, and real-world scenarios would involve specific data and parameters related to the diagnostic test and the disease in question.

Q7. Calculate the 95% confidence interval for a sample of data with a mean of 50 and a standard deviation

of 5. Interpret the results.

Ans:To calculate the 95% confidence interval for a sample mean, we can use the formula:

Confidence Interval=(�¯−���,�¯+���)

Confidence Interval=(

x

$-$

$-Z$

n

s

,

x

$-$

$+Z$

n

s

)

Where:

- �‾
  - x
  - ‾
- is the sample mean.
- �
- s is the sample standard deviation.
- �
- n is the sample size.
- �
- Z is the Z-score corresponding to the desired confidence level.

For a 95% confidence interval, the Z-score is approximately 1.96.

Given the information:

- �‾=50
  - x
  - ‾
- $=50$ (sample mean)
- �=5
- $s=5$ (sample standard deviation)

Let's calculate the confidence interval:

Confidence Interval=(50−1.965�,50+1.965�)

Confidence Interval=(50−1.96

n

5

,50+1.96

n

5

)

However, to calculate the confidence interval, we need the sample size (

�

n). If you have a specific sample size, you can substitute it into the formula.

Assuming a sample size of 100 for the purpose of illustration:

python

Copy code

```
import
```

```
50
5
100
0.95
1.96
print f"95% Confidence Interval: ({lower_bound:.2f}, {upper_bound:.2f})"
```

Interpretation:

- We are 95% confident that the true population mean falls within the interval (lower_bound, upper_bound).
- In this case, if we were to repeat this process for many random samples, 95% of the calculated intervals would contain the true population mean.

Q8. What is the margin of error in a confidence interval? How does sample size affect the margin of error?

Provide an example of a scenario where a larger sample size would result in a smaller margin of error.

Ans:The margin of error in a confidence interval is a measure of the uncertainty or precision associated with the estimate of a population parameter (such as the mean or proportion) based on a sample. It represents the range above and below the point estimate within which we are reasonably confident the true parameter value lies.

The general formula for the margin of error in a confidence interval for the population mean (

�

μ) is given by:

Margin of Error=���

Margin of Error=Z

n

s

Where:

- �
- Z is the Z-score corresponding to the desired confidence level.
- �
- s is the sample standard deviation.
- �
- n is the sample size.

The margin of error is influenced by the following factors:

Confidence Level (
�
Z): A higher confidence level requires a larger Z-score, resulting in a wider interval and a larger margin of error.

Sample Standard Deviation (
�
s): A larger standard deviation increases the variability in the data, leading to a wider interval and a larger margin of error.

Sample Size (
�
n): The most notable factor affecting the margin of error is the sample size. As the sample size increases, the margin of error decreases. This is because larger sample sizes provide more information about the population, leading to a more precise estimate.

Example Scenario:

Suppose you want to estimate the average income of households in a city. You take two samples, one with a small sample size (e.g., 50 households) and another with a larger sample size (e.g., 500 households). Assuming the same confidence level, standard deviation, and point estimate, the margin of error will be smaller for the larger sample size.

python

Copy code

```
import

                    0.95

                      1000

                  50000

                     50

          1.96

                     500
```

```
print f"Margin of Error (Small Sample): {margin_of_error_small:.2f}"

print f"Margin of Error (Large Sample): {margin_of_error_large:.2f}"
```

In this example, you'll observe that the margin of error is smaller for the larger sample size, indicating increased precision in estimating the average income. This demonstrates the inverse relationship between sample size and the margin of error: larger samples result in smaller margins of error, providing more confidence in the precision of the estimate.

Q9. Calculate the z-score for a data point with a value of 75, a population mean of 70, and a population

standard deviation of 5. Interpret the results.

Ans:The Z-score, also known as the standard score or z-value, measures how many standard deviations a data point is from the mean of a distribution. It is calculated using the formula:

�=(�−�)�

$Z=$

$$\sigma$$

$$(X-\mu)$$

Where:

- �
- Z is the Z-score.
- �
- X is the individual data point.
- �
- μ is the population mean.
- �
- σ is the population standard deviation.

In this case, the values are:

- �=75
- $X=75$ (data point)
- �=70
- $\mu=70$ (population mean)
- �=5
- $\sigma=5$ (population standard deviation)

Let's calculate the Z-score:

�=(75−70)5

$Z=$

$$\frac{(75-70)}{5}$$

python

Copy code

```
75

70

5

print f"The Z-score for the data point is: {z_score:.2f}"
```

Interpretation:

- A Z-score of 1 corresponds to being 1 standard deviation above the mean. Therefore, a Z-score of
- �=1
- $Z=1$ means the data point 75 is 1 standard deviation above the population mean of 70.
- A positive Z-score indicates that the data point is above the mean, while a negative Z-score would indicate that the data point is below the mean.
- In this case, the value of 75 is higher than the mean, and the Z-score quantifies how many standard deviations above the mean it is.

Q10. In a study of the effectiveness of a new weight loss drug, a sample of 50 participants lost an average of 6 pounds with a standard deviation of 2.5 pounds. Conduct a hypothesis test to determine if the drug is

significantly effective at a 95% confidence level using a t-test.

Q10. In a study of the effectiveness of a new weight loss drug, a sample of 50 participants lost an average

of 6 pounds with a standard deviation of 2.5 pounds. Conduct a hypothesis test to determine if the drug is

significantly effective at a 95% confidence level using a t-test.

Ans:To conduct a hypothesis test for the effectiveness of the weight loss drug, we can use a t-test for the population mean. The null hypothesis (

�0

H

0

) is that the average weight loss (

�

μ) is zero or not significantly different from zero, and the alternative hypothesis (

�1

H

1

or

��

H

a

) is that the average weight loss is significantly different from zero.

The general form of the hypotheses is as follows:

�0:�=0

H

0

$:\mu=0$

�1:�≠0

H

1

$:\mu$

$=0$

The formula for the t-statistic in this case is:

�=�‾−�0��

$t=$

n

s

x

-

$-\mu$

0

Where:

- �‾
- x
- ‾
- is the sample mean (average weight loss).
- �0
- μ
- 0
-
- is the hypothesized population mean under the null hypothesis (usually 0 in this case).
- �
- s is the sample standard deviation.
- �
- n is the sample size.

Given the information:

- Sample mean (
- �̄
  - x
  - ̄
-) = 6 pounds
- Sample standard deviation (
- �
- s) = 2.5 pounds
- Sample size (
- �
- n) = 50
- Hypothesized population mean under the null hypothesis (
- �0
- μ
- 0
-
-) = 0
- Significance level = 0.05 (for a 95% confidence level)

Let's perform the calculations:

python

Copy code

```
import

from            import




              6
```

2.5

50

0

0.95

1

1 1 2

2 1 abs

```
print f"t-statistic: {t_statistic}"

print f"Critical value: {critical_value}"

print f"P-value: {p_value}"



if abs

 print "Reject the null hypothesis. The drug is significantly effective."

else

 print "Fail to reject the null hypothesis. There is not enough evidence to
conclude that the drug is significantly effective."
```

Interpretation:

- The t-statistic measures how many standard errors the sample mean is from the hypothesized population mean.
- The critical value is determined based on the significance level (0.05 for a 95% confidence level).
- The p-value is the probability of observing a t-statistic as extreme as the one calculated.
- If the absolute value of the t-statistic is greater than the critical value, or if the p-value is less than the significance level, we reject the null hypothesis.

Please note that the choice of the significance level (0.05 in this case) is a common convention, but it can be adjusted based on the specific requirements of the study.

Q11. In a survey of 500 people, 65% reported being satisfied with their current job. Calculate the 95%

confidence interval for the true proportion of people who are satisfied with their job.

Ans:To calculate the 95% confidence interval for the true proportion of people who are satisfied with their job, we can use the formula for the confidence interval for a population proportion. The formula is given by:

Confidence Interval=(�^−��^(1−�^)�,�^+��^(1−�^)�)

Confidence Interval=(

p

^

$-Z$

n

p

^

(1−

p

^

)

,

p

^

$+Z$

n

p

^

(1−

p

^

)

)

Where:

- �^
  - p
  - ^
  -
- is the sample proportion (percentage converted to a decimal).

- �
- Z is the Z-score corresponding to the desired confidence level.
- �
- n is the sample size.

Given the information:

- Sample proportion (
- �^
  - p
  - ^
  -
-) = 0.65 (65% expressed as a decimal)
- Sample size (
- �
- n) = 500
- Confidence level = 0.95

Let's calculate the confidence interval:

python

Copy code

```
import

0.65

500
```

0.95

1.96

1

```
print f"95% Confidence Interval for the Proportion: ({lower_bound:.4f},
{upper_bound:.4f})"
```

Interpretation:

- We are 95% confident that the true proportion of people satisfied with their job falls within the interval (lower_bound, upper_bound).
- In this case, if we were to repeat this process for many random samples, 95% of the calculated intervals would contain the true population proportion.

Q12. A researcher is testing the effectiveness of two different teaching methods on student performance.

Sample A has a mean score of 85 with a standard deviation of 6, while sample B has a mean score of 82

with a standard deviation of 5. Conduct a hypothesis test to determine if the two teaching methods have a

significant difference in student performance using a t-test with a significance level of 0.01.

Ans:To conduct a hypothesis test to determine if there is a significant difference in student performance between two teaching methods, we can use a two-sample t-test. The null hypothesis (

�0

H

0

) is that there is no significant difference in means, and the alternative hypothesis (

�1

H

1

or

��

H

a

) is that there is a significant difference.

The general form of the hypotheses is as follows:

�0:�1=�2

H

0

$:\mu$

1

$=\mu$

2

�1:�1≠�2

H

1

$:\mu$

1

$=\mu$

2

The formula for the t-statistic in this case is:

�=(�¯1−�¯2)�12�1+�22�2

$t=$

n

1

s

1

2

$+$

n

2

s

2

2

(

x

-

1

–

x

-

2

)

Where:

- �¯1
- x
- ¯
- 1
-
- and
- �¯2
- x
- ¯
- 2
-
- are the sample means for Sample A and Sample B, respectively.
- �1
- s

- 1
-
- and
- �2
- s
- 2
-
- are the sample standard deviations for Sample A and Sample B, respectively.
- �1
- n
- 1
-
- and
- �2
- n
- 2
-
- are the sample sizes for Sample A and Sample B, respectively.

Given the information:

- Sample A:
- �¯1=85
  - x
  - ¯
- 1
-
- =85,
- �1=6
- s
- 1
-
- =6,
- �1=?
- n
- 1

-
- =? (not provided)
- Sample B:
- �¯2=82
- x
- ¯
- 2
-
- =82,
- �2=5
- s
- 2
-
- =5,
- �2=?
- n
- 2
-
- =? (not provided)
- Significance level = 0.01

Since the sample sizes (

�1

n

1

and

�2

n

2

) are not provided, I'll assume equal sample sizes for simplicity. You may adjust the values based on the actual sizes.

Let's perform the calculations:

python

Copy code

```
import

85

6

82

5

0.01

30

2

2
```

```
                  2                  2

               abs                             2

          2    1         abs

print f"t-statistic: {t_statistic}"

print f"Critical value: {critical_value}"

print f"P-value: {p_value}"

if abs
```

```
 print "Reject the null hypothesis. There is a significant difference in student
performance between the two teaching methods."

else

 print "Fail to reject the null hypothesis. There is not enough evidence to
conclude a significant difference in student performance."
```

Note: The sample sizes are assumed to be equal for simplicity. Adjust the sample sizes based on the actual data if available. Additionally, the assumption of equal variances between the two samples is made here. If this assumption is not valid, a modified version of the t-test (e.g., Welch's t-test) may be more appropriate.

Q13. A population has a mean of 60 and a standard deviation of 8. A sample of 50 observations has a mean

of 65. Calculate the 90% confidence interval for the true population mean.

Ans:To calculate the 90% confidence interval for the true population mean, we can use the formula for the confidence interval for a population mean. The formula is given by:

Confidence Interval=(�¯−���,�¯+���)

Confidence Interval=(

x

-

$-Z$

n

σ

,

x

$-$

$+Z$

n

σ

)

Where:

- �$^{-}$
- x
- $^{-}$
- is the sample mean.
- �
- Z is the Z-score corresponding to the desired confidence level.
- �
- σ is the population standard deviation.
- �
- n is the sample size.

Given the information:

- Sample mean (
- �‾

- x
- ‾

-) = 65
- Population standard deviation (
- �
- σ) = 8
- Sample size (
- �
- n) = 50
- Confidence level = 0.90

Let's calculate the confidence interval:

python

Copy code

```
import

65

8

50

0.90
```

```
1.645
```

```
print f"90% Confidence Interval for the Population Mean: ({lower_bound:.2f},
{upper_bound:.2f})"
```

Interpretation:

- We are 90% confident that the true population mean falls within the interval (lower_bound, upper_bound).
- In this case, if we were to repeat this process for many random samples, 90% of the calculated intervals would contain the true population mean.

Q14. In a study of the effects of caffeine on reaction time, a sample of 30 participants had an average

reaction time of 0.25 seconds with a standard deviation of 0.05 seconds. Conduct a hypothesis test to

determine if the caffeine has a significant effect on reaction time at a 90% confidence level using a t-test.

Ans:To conduct a hypothesis test to determine if caffeine has a significant effect on reaction time, we can use a one-sample t-test. The null hypothesis (

�0

H

0

) is that there is no significant effect (the mean reaction time is equal to a specified value), and the alternative hypothesis (

�1

H

1

or

��

H

a

) is that there is a significant effect.

The general form of the hypotheses is as follows:

�0:�=�0

H

0

$:\mu=\mu$

0

�1:�≠�0

H

1

$:\mu$

$=\mu$

0

The formula for the t-statistic in this case is:

�=(�¯−�0)��

$t=$

n

s

(

x

‾

$-\mu$

0

)

Where:

- �‾
    - x
    - ‾
- is the sample mean (average reaction time).
- �0
- μ
- 0
-
- is the hypothesized population mean under the null hypothesis.
- �
- s is the sample standard deviation.
- �
- n is the sample size.

Given the information:

- Sample mean (
- �‾
- x
- ‾
-) = 0.25 seconds
- Sample standard deviation (
- �
- s) = 0.05 seconds
- Sample size (
- �
- n) = 30
- Hypothesized population mean under the null hypothesis (
- �0
- μ
- 0
-
-) = ? (not provided)
- Confidence level = 0.90

For the purpose of this example, let's assume that the null hypothesis is that caffeine has no effect on reaction time, so

�0=0

μ

0

=0 (zero effect).

Let's perform the calculations:

python

Copy code

```
import
0.25
0.05
30
0
0.90
1
```

1 1 2

2 1 abs

```
print f"t-statistic: {t_statistic}"

print f"Critical value: {critical_value}"

print f"P-value: {p_value}"

if abs

 print "Reject the null hypothesis. Caffeine has a significant effect on reaction time."

else

 print "Fail to reject the null hypothesis. There is not enough evidence to conclude a significant effect of caffeine on reaction time."
```

Note: The choice of the null hypothesis depends on the specific question being addressed in the study. Adjust the null hypothesis based on the research question and context.