Here is a structured answer to your assignment questions:

**Q1: What is Estimation Statistics? Explain point estimate and interval estimate.**

**Estimation Statistics** involves methods to estimate population parameters based on sample data.

* **Point Estimate:** A single value estimate of a population parameter (e.g., sample mean for population mean).
* **Interval Estimate:** A range of values within which the population parameter is expected to lie, often expressed with a confidence level (e.g., 95% confidence interval).

**Q2: Python function to estimate population mean.**

def estimate\_population\_mean(sample\_mean, sample\_std, sample\_size, confidence=0.95):

from scipy.stats import t

import math

alpha = 1 - confidence

t\_critical = t.ppf(1 - alpha / 2, df=sample\_size - 1)

margin\_of\_error = t\_critical \* (sample\_std / math.sqrt(sample\_size))

lower\_bound = sample\_mean - margin\_of\_error

upper\_bound = sample\_mean + margin\_of\_error

return lower\_bound, upper\_bound

**Q3: What is Hypothesis testing? Why is it used? Importance.**

**Hypothesis Testing** is a statistical method to determine if there is enough evidence to support a specific claim about a population parameter.

* **Uses:** To validate assumptions, compare groups, or test relationships.
* **Importance:** Reduces uncertainty, guides decision-making, and ensures objective analysis.

**Q4: Hypothesis: Average weight comparison.**

* **Null Hypothesis (H₀):** The average weight of male college students is equal to or less than the average weight of female college students.
* **Alternative Hypothesis (H₁):** The average weight of male college students is greater than the average weight of female college students.

**Q5: Python script for hypothesis test (two population means).**

from scipy.stats import ttest\_ind

def hypothesis\_test\_two\_means(sample1, sample2, alpha=0.05):

t\_stat, p\_value = ttest\_ind(sample1, sample2)

if p\_value < alpha:

return "Reject the null hypothesis"

else:

return "Fail to reject the null hypothesis"

**Q6: Null and Alternative Hypothesis.**

* **Null Hypothesis (H₀):** Assumes no effect or difference (e.g., population means are equal).
* **Alternative Hypothesis (H₁):** Assumes an effect or difference exists (e.g., population means are not equal).

**Examples:**

1. H₀: The new drug has no effect. H₁: The new drug reduces symptoms.
2. H₀: Mean weight = 5 kg. H₁: Mean weight ≠ 5 kg.

**Q7: Steps in Hypothesis Testing.**

1. State the null and alternative hypotheses.
2. Choose a significance level (α).
3. Collect data and compute a test statistic.
4. Determine the critical value or p-value.
5. Make a decision: reject or fail to reject the null hypothesis.

**Q8: Define p-value and significance.**

* **p-value:** The probability of obtaining a test statistic at least as extreme as the observed value under H₀.
* **Significance:** A small p-value (< α) indicates strong evidence against H₀.

**Q9: Generate Student's t-distribution plot.**

import numpy as np

import matplotlib.pyplot as plt

from scipy.stats import t

x = np.linspace(-4, 4, 1000)

y = t.pdf(x, df=10)

plt.plot(x, y, label="t-distribution (df=10)")

plt.title("Student's t-Distribution")

plt.xlabel("x")

plt.ylabel("Density")

plt.legend()

plt.show()

**Q10: Python program for two-sample t-test.**

from scipy.stats import ttest\_ind

def two\_sample\_t\_test(sample1, sample2, alpha=0.05):

t\_stat, p\_value = ttest\_ind(sample1, sample2)

return "Reject H₀" if p\_value < alpha else "Fail to reject H₀"

**Q11: Student’s t-Distribution.**

* **Definition:** A probability distribution used when the sample size is small, and population standard deviation is unknown.
* **Use:** Small sample sizes or when standard deviation is estimated from the sample.

**Q12: t-statistic formula.**

t=xˉ−μs/nt = \frac{\bar{x} - \mu}{s / \sqrt{n}}

Where:

* xˉ\bar{x}: Sample mean
* μ\mu: Population mean
* ss: Sample standard deviation
* nn: Sample size

**Q13: Estimate population mean with 95% confidence interval.**

Given:

* Sample mean = $500
* Standard deviation = $50
* Sample size = 50
* Confidence level = 95%

Solution (using Python):

sample\_mean = 500

sample\_std = 50

sample\_size = 50

ci = estimate\_population\_mean(sample\_mean, sample\_std, sample\_size)

print(f"95% confidence interval: {ci}")

**Q14: Hypothesis test for drug effect.**

sample\_mean = 8

pop\_mean = 10

std\_dev = 3

n = 100

alpha = 0.05

t\_stat = (sample\_mean - pop\_mean) / (std\_dev / np.sqrt(n))

critical\_value = t.ppf(1 - alpha, df=n - 1)

result = "Reject H₀" if t\_stat < -critical\_value else "Fail to reject H₀"

print(result)

**Q15: Hypothesis test for mean weight.**

Given:

* Sample mean = 4.8, Population mean = 5, Sample std = 0.5, Sample size = 25, α\alpha = 0.01

Python Solution:

t\_stat = (4.8 - 5) / (0.5 / np.sqrt(25))

critical\_value = t.ppf(0.01, df=24)

result = "Reject H₀" if t\_stat < critical\_value else "Fail to reject H₀"

print(result)

**Q16: Hypothesis test for group means.**

n1, mean1, std1 = 30, 80, 10

n2, mean2, std2 = 40, 75, 8

alpha = 0.01

pooled\_std = np.sqrt(((n1 - 1) \* std1\*\*2 + (n2 - 1) \* std2\*\*2) / (n1 + n2 - 2))

t\_stat = (mean1 - mean2) / (pooled\_std \* np.sqrt(1/n1 + 1/n2))

critical\_value = t.ppf(1 - alpha / 2, df=n1 + n2 - 2)

result = "Reject H₀" if abs(t\_stat) > critical\_value else "Fail to reject H₀"

print(result)

**Q17: Estimation for number of ads watched.**

Given:

* Sample mean = xˉ\bar{x}, sample size = 50, confidence = 95%.

Solution (similar to Q13).