Introduction to Statistics - Quiz #3(60 minutes)

December 2, 2025 (Tuesday)

Section(교반): _____ Cadet Number(교번): ____________ Name(성명): __________________ Score: ____________

- All solutions must include a detailed step-by-step explanation.
- If an answer has more than four decimal places, round to the **fourth decimal place**.
- Reference table is provided on the last page of the exam.

1. A factory wants to estimate the mean operating time of machines μ. A random sample of $n = 15$ machines was selected, and their operating times (in minutes) were measured as follows:

i	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
x_i	304.71	293.02	296.60	301.85	294.92	299.64	300.90	295.84	293.45	300.97	304.97	296.77	298.33	308.23	297.21

The sample mean is $\bar{x} = 299.16$, and the sample standard deviation is $s = 4.44$. Assuming that all the necessary conditions are satisfied, construct a **90%** confidence interval for the mean operating time μ. [20 points]

Solution: To construct a 90% confidence interval for mean μ, we use the t-distribution. The degrees of freedom are $df = n - 1 = 14$. Therefore, the 90% confidence interval is:

$$\bar{x} \pm t_{0.05}(14) \cdot \frac{s}{\sqrt{n}} = 299.16 \pm 1.7613 \cdot \frac{4.44}{\sqrt{15}}$$
$$\approx 299.16 \pm 2.0192 = (297.1408, 301.1792)$$

2. A company reports that the average strength of its steel bars, denoted by μ, is 655 MPa. Recently, an expert claimed that the average strength might be **lower** than 655 MPa. To check this claim, a random sample of $n = 9$ steel bars was selected, and their strengths (in MPa) were measured. We conduct **a hypothesis test for a population mean** at the significance level $\alpha = 0.05$. (Assume that all the necessary conditions are satisfied.) [40 points]

(1) State the null and alternative hypotheses. (Use a **one-sided** test.)

Solution:

$$H_0 : \mu = 655, \qquad H_A : \mu < 655$$

(2) Find the test statistic and its null distribuiton.

Solution: Under H_0, the test statistic follows:

$$T = \frac{\bar{X} - \mu_0}{S/\sqrt{n}} = \frac{\bar{X} - 655}{S/\sqrt{9}} \overset{H_0}{\sim} t(8)$$

where $\bar{X}$ is the sample mean, and S is the sample standard deviation.

(3) The sample mean is $\bar{x} = 648$, and the sample standard deviation is $s = 21$. Compute the observed test statistic.

Solution:

$$t = \frac{648 - 655}{21/\sqrt{9}} = \frac{-7}{7} = \boxed{-1}$$

(4) Report the p-value and complete the hypothesis test. State the conclusion in the context of the data.

Solution:

$$p\text{-value} = P(T < -1) = 0.1733, \text{ where } T \sim t(8).$$

Since the p-value is greater than $\alpha = 0.05$, we fail to reject H_0.

Conclusion: At the significance level $\alpha = 0.05$, there is insufficient evidence to conclude that the true mean strength of the company's steel bars is less than 655 MPa.

3. A researcher wants to know whether adults who drink soda every day consume different amounts of sugar compared to adults who rarely drink soda. Daily sugar intake (grams per day) was recorded for $n_E = 12$ everyday soda drinkers and $n_R = 10$ rare soda drinkers. We conduct a **two-sample t-test** at the significance level $\alpha = 0.05$ to determine whether the population mean of daily sugar intake for everyday soda drinkers (μ_E) differs from that of rare soda drinkers (μ_R). **Assume that the variances of daily sugar intake are equal for both groups** ($\sigma_E^2 = \sigma_R^2 = \sigma^2$). [40 points]

(1) State the null and alternative hypotheses. (Use a **two-sided** test.)

Solution:

$$H_0 : \mu_E - \mu_R = 0, \qquad H_A : \mu_E - \mu_R \neq 0$$

(2) Find the test statistic and its null distribution.

Solution: Since we assume equal variances, we use the pooled standard deviation. Under H_0, the test statistic is:

$$T = \frac{\bar{X}_E - \bar{X}_R}{S_p\sqrt{\frac{1}{n_E} + \frac{1}{n_R}}} = \frac{\bar{X}_E - \bar{X}_R}{S_p \cdot \sqrt{\frac{1}{12} + \frac{1}{10}}} \overset{H_0}{\sim} t(df)$$

where S_p is the pooled standard deviation and $df = n_E + n_R - 2 = 20$.

(3) Daily sugar intake was measured for 12 everyday soda drinkers and 10 rare soda drinkers. Using the R code and partial R output provided, report the p-value and complete the hypothesis test. State the conclusion in the context of the data.

```
everyday = c(64, 70, 59, 68, 61, 60, 66, 72, 58, 63, 69, 65); rare = c(58, 65, 55, 62, 59, 56, 60, 57, 61, 54)
t.test(everyday, rare, alternative = "two.sided", mu = 0, var.equal = TRUE)
```

```
        Two Sample t-test
data:  everyday and rare
t = 3.3673, df = [    ], p-value = [    ]
...(omitted)...
```

Solution: The observed test statistic is $t \approx 3.3673$, and the p-value is $2 \times P(T > 3.3673) \approx 0.0030$, where $T \sim t(20)$. Since the p-value is less than $\alpha = 0.05$, we reject the null hypothesis.

Conclusion: There is sufficient statistical evidence to conclude that the mean daily sugar intake differs between adults who drink soda every day and those who rarely drink soda.

(4) Let $\bar{X}_E$ and $\bar{X}_R$ be the sample means for the two groups. Suppose a random variable V satisfies the following:

$$\frac{\bar{X}_E - \bar{X}_R - (\mu_E - \mu_R)}{S_p\sqrt{\frac{1}{n_E} + \frac{1}{n_R}}} = \frac{\bar{X}_E - \bar{X}_R - (\mu_E - \mu_R)}{\sigma\sqrt{\frac{1}{n_E} + \frac{1}{n_R}}} \times \left(\frac{V}{n_E + n_R - 2}\right)^{-1/2}$$

where S_p is the pooled sample standard deviation. Simplify V and find the distribution of V.

Solution:

$$V = \frac{(n_E + n_R - 2)S_p^2}{\sigma^2} \sim \chi^2(n_E + n_R - 2).$$

Reference Table

Satterthwaite's df : $\psi = \frac{(s_1^2/n_1 + s_2^2/n_2)^2}{(s_1^2/n_1)^2/(n_1-1) + (s_2^2/n_2)^2/(n_2-1)}$, $S_p^2 = \frac{(n_1-1)S_1^2 + (n_2-1)S_2^2}{n_1+n_2-2}$

$z_{0.005} = 2.5758$	$z_{0.01} = 2.3263$	$z_{0.025} = 1.9600$	$z_{0.05} = 1.6449$
$t_{0.025}(14) = 2.1448$	$t_{0.05}(14) = 1.7613$	$t_{0.025}(15) = 2.1315$	$t_{0.05}(15) = 1.7531$

$P(T < -1.0000) = 0.1733, \quad T \sim t(8)$	$P(T < -1.0000) = 0.1717, \quad T \sim t(9)$
$P(T < -0.3333) = 0.3737, \quad T \sim t(8)$	$P(T < -0.3333) = 0.3733, \quad T \sim t(9)$
$P(T > 3.3673) = 0.0015, \quad T \sim t(20)$	$P(T > 3.3673) = 0.0036, \quad T \sim t(10)$
$P(T > 1.6837) = 0.0539, \quad T \sim t(20)$	$P(T > 1.6837) = 0.0616, \quad T \sim t(10)$