Day 53 Notes: Unit 10.2 Two Sample Means T Interval **AP Statistics**

Name ____________________________________ **Block** _______ **Date** ____________

Shape: Both samples must be approximately Normal by any of the ff:

- Population is approximately normal.
- $n \geq 30$ Central Limit Theorem
- If $n < 30$, the graph shows no strong skewness or outliers.

Center:
$\mu_{\bar{X}_1-\bar{X}_2} = \mu_1 - \mu_2$

Spread:

$$\sigma_{\bar{X}_1-\bar{X}_2} = \sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}$$

Two Sample Means T Interval
Condition:

- SRS, 10% Condition, Normal

$$\bar{X}_1 - \bar{X}_2 \pm t^* \sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}$$

$$d.f. \text{ is smaller of } n_1 - 1 \text{ or } n_2 - 1.$$

Check Your Understanding

1. The most recent American Time Use Survey, conducted by the Bureau of Labor Statistics, found that many Americans barely spend any time reading for fun. People ages 15 to 19 average only 7.8 minutes of leisurely reading per day with a standard deviation of 5.4 minutes. However, people ages 75 and over read for an average of 43.8 minutes per day with a standard deviation of 35.5 minutes. These results were based on random samples of 975 people ages 15 to 19 and 1050 people ages 75 and over.

Hypotheses:

Null (H0): Mean reading time for ages 15-19 equals mean for ages 75 and over.
Alternative (Ha): Means differ (two-tailed test).
Test Statistic:

t = (mean1 - mean2) / sqrt((s1^2 / n1) + (s2^2 / n2))

mean1 = 7.8, s1 = 5.4, n1 = 975
mean2 = 43.8, s2 = 35.5, n2 = 1050
Calculate standard error:

(5.4^2 / 975) = 29.16 / 975 ≈ 0.0299

(35.5^2 / 1050) = 1260.25 / 1050 ≈ 1.2002

sqrt(0.0299 + 1.2002) = sqrt(1.2301) ≈ 1.109

t = (7.8 - 43.8) / 1.109 = -36 / 1.109 ≈ -32.46

Degrees of Freedom: With large samples (975 and 1050), degrees of freedom are large (over 1000), so the t-distribution approximates the normal distribution. The critical t-value for alpha = 0.05 (two-tailed) is about 1.96. Since |t| = 32.46 > 1.96, we reject H0.

P-value: A t-statistic of -32.46 yields a p-value far less than 0.05 (essentially 0), confirming statistical significance.

95% Confidence Interval for mean1 - mean2:

(mean1 - mean2) ± t* × sqrt((s1^2 / n1) + (s2^2 / n2))

-36 ± 1.96 × 1.109

1.96 × 1.109 ≈ 2.174

-36 - 2.174 = -38.174, -36 + 2.174 = -33.826

Interval: (-38.17, -33.83) minutes. Since it's all negative and excludes 0, the means differ significantly.

Conclusion: The t-statistic of -32.46 and p-value < 0.05 show a significant difference. People ages 75 and over read, on average, 36 minutes more per day than those ages 15-19, with a 95% confidence interval of 33.83 to 38.17 minutes more (reversing signs: 43.8 - 7.8 = 36, interval 33.83 to 38.17). Thus, older Americans spend significantly more time reading for fun than younger ones.

2. Mr. Wilcox's class performed an experiment to investigate whether drinking a caffeinated beverage would increase pulse rates. Twenty students in the class volunteered to take part in the experiment. All of the students measured their initial pulse rates (in beats per minute). Then Mr. Wilcox randomly assigned the students into two groups of 10. Each student in the first group drank 12 ounces of cola with caffeine. Each student in the second group drank 12 ounces of caffeine-free cola. All students then measured their pulse rates again. The table displays the change in pulse rate for the students in both groups.

	Change in pulse rate (Final pulse rate − Initial pulse rate)										Mean change
Caffeine	8	3	5	1	4	0	6	1	4	0	3.2
No caffeine	3	−2	4	−1	5	5	1	2	−1	4	2.0

Construct and interpret a 95% confidence interval for the difference in true mean change in pulse rate for subjects like these who drink caffeine versus who drink no caffeine. Interpret the confidence interval and confidence level.

Caffeine variance (s1^2):

Sum of squared deviations = 65.6 (calculated from data)

s1^2 = 65.6 / 9 ≈ 7.29

No-caffeine variance (s2^2):

Sum of squared deviations = 62 (calculated from data)

s2^2 = 62 / 9 ≈ 6.89

Since variances are close (7.29 / 6.89 ≈ 1.06), assume equal variances:

sp^2 = (65.6 + 62) / (10 + 10 - 2) = 127.6 / 18 ≈ 7.09

Standard error:

SE = square root of [(7.09 / 10) + (7.09 / 10)] = square root of 1.418 ≈ 1.19

df = 10 + 10 - 2 = 18

t* for 95% confidence (df = 18) ≈ 2.1 (from t-table)

Difference: 3.2 - 2.0 = 1.2

Margin of error = 2.1 × 1.19 ≈ 2.5

Interval: 1.2 - 2.5 = -1.3, 1.2 + 2.5 = 3.7

95% CI: (-1.3, 3.7) beats per minute

The 95% confidence interval for the difference in true mean change (caffeine minus no-caffeine) is (-1.3, 3.7) beats per minute. Since it includes zero, we can’t conclude a significant difference in pulse rate change between groups. The 95% confidence level means 95% of such intervals from repeated samples would contain the true difference.