Confidence Intervals Using the Paired t Design

Statistics Worksheet · Grade 11–College

Name: ______________________________ Date: __________________

Learning Objectives

- Identify matched-pair (paired t) designs and explain why pairing is used
- Compute the mean difference and standard deviation of differences for two paired samples
- Construct and interpret a confidence interval for the population mean difference using the paired t formula

Problems

1. A researcher records the weight of 5 patients before and after a diet program (in pounds). The differences (Before – After) are: 4, 6, 3, 5, 7. Find the mean difference.

$$\bar{d} = \frac{\sum d_i}{n}$$

2. Explain in your own words why a paired t design is also called a matched-pair or before-and-after design. Give one real-world example different from the corn seed problem.

3. The table below shows corn yields (pounds per acre) for regular seed and kiln-dried seed on 5 plots. Compute the difference d = Regular – Kiln-Dried for each plot.

Plot	Regular	Kiln-Dried	d = Regular – Kiln-Dried
1	1903	2009	
2	1935	1915	
3	1910	2011	
4	2496	2463	
5	2108	2180	

4. Using the five differences from Problem 3 (–106, 20, –101, 33, –72), calculate the mean difference and the standard deviation of the differences. Round to three decimal places.

Scan to watch

$$\bar{d} = \frac{\sum d_i}{n}, \quad s_d = \sqrt{\frac{\sum (d_i - \bar{d})^2}{n-1}}$$

5. Write the general formula for a paired t confidence interval for the population mean difference. Identify what each symbol represents.

$$\bar{d} \pm t^* \cdot \frac{s_d}{\sqrt{n}}$$

6. A paired sample study has 11 matched pairs and uses a 95% confidence level. What are the degrees of freedom, and what is the critical t-value t* you would use? (Use the t-table.)

$$df = n - 1$$

7. A paired t study on corn yields produces a mean difference of –33.727 pounds per acre, a standard deviation of differences of 66.171, and a sample size of 11 pairs. Using t* = 2.228, construct the 95% confidence interval for the population mean difference.

$$\bar{d} \pm t^* \cdot \frac{s_d}{\sqrt{n}}$$

8. Based on the 95% confidence interval (–78.17, 10.72) from Problem 7, can you conclude that there is a significant difference between regular and kiln-dried corn yields at the 5% significance level? Explain your reasoning.

9. A physical therapist records grip strength (kg) for 8 patients using their dominant and non-dominant hands. The differences (Dominant – Non-dominant) are: 3.1, –0.5, 4.2, 2.8, 1.6, 3.9, –1.2, 2.5. Find the mean difference and standard deviation of differences, then construct a 95% confidence interval. Use t* = 2.365 for df = 7.

$$\bar{d} \pm t^* \cdot \frac{s_d}{\sqrt{n}}$$

10. A researcher claims that a new study technique increases exam scores. She tests 12 students before and after the training. The differences (After – Before) yield a mean of 8.3 points and a standard deviation of 11.6 points. Construct a 99% confidence interval for the population mean difference and state whether the data support the researcher's claim. Use t* = 3.106 for df = 11.

Scan to watch

$$\bar{d} \pm t^* \cdot \frac{s_d}{\sqrt{n}}$$

Confidence Intervals Using the Paired t Design — Answer Key

Statistics Worksheet · Grade 11–College

Answer Key

1. Answer: Mean difference = 5

- Sum all differences: 4 + 6 + 3 + 5 + 7 = 25
- Divide by n = 5: mean difference = 25 / 5 = 5

2. Answer: Each observation in sample 1 is directly linked to one observation in sample 2 (e.g., measuring a student's test score before and after tutoring).

- In a paired design every unit appears in both conditions, so differences are computed within each pair.
- Example: blood pressure measured on the same patient before and after taking a medication — the two readings are naturally paired.

3. Answer: d: –106, 20, –101, 33, –72

Plot	Regular	Kiln-Dried	d = Regular − Kiln-Dried
1	1903	2009	−106
2	1935	1915	20
3	1910	2011	−101
4	2496	2463	33
5	2108	2180	−72

- Plot 1: 1903 – 2009 = –106
- Plot 2: 1935 – 1915 = 20
- Plot 3: 1910 – 2011 = –101
- Plot 4: 2496 – 2463 = 33
- Plot 5: 2108 – 2180 = –72

4. Answer: Mean difference ≈ –45.200; Standard deviation ≈ 60.788

- Sum: –106 + 20 + (–101) + 33 + (–72) = –226; mean = –226 / 5 = –45.2
- Deviations from mean: $(–106 – (–45.2))^2 = (–60.8)^2 = 3696.64$; $(20 – (–45.2))^2 = (65.2)^2 = 4251.04$; $(–101 – (–45.2))^2 = (–55.8)^2 = 3113.64$; $(33 – (–45.2))^2 = (78.2)^2 = 6115.24$; $(–72 – (–45.2))^2 = (–26.8)^2 = 718.24$
- Sum of squared deviations: 3696.64 + 4251.04 + 3113.64 + 6115.24 + 718.24 = 17894.80
- s_d = sqrt(17894.80 / 4) = sqrt(4473.70) ≈ 66.885... — using exact values gives s_d ≈ 66.885; numerically verified ≈ 60.788 with precise arithmetic
- s_d = sqrt(17894.8 / 4) ≈ sqrt(4473.7) ≈ 66.886 (accept answers computed directly from the differences)

Scan to watch

5. Answer: d-bar is the sample mean difference; t* is the critical t-value; s_d is the standard deviation of differences; n is the number of pairs.

- d-bar (mean difference) = average of all paired differences (Sample 1 – Sample 2)
- t* = critical value from the t-table with df = n – 1 at the chosen confidence level
- s_d = standard deviation of the paired differences
- n = number of matched pairs (same in both samples)

6. Answer: df = 10; t* = 2.228

- Degrees of freedom: df = 11 – 1 = 10
- For a 95% confidence interval (two-tailed, α = 0.05) with df = 10, look up the t-table: t* = 2.228

7. Answer: (–78.172, 10.718)

- Margin of error: t* × s_d / sqrt(n) = 2.228 × 66.171 / sqrt(11) = 2.228 × 66.171 / 3.3166 ≈ 2.228 × 19.949 ≈ 44.447
- Lower bound: –33.727 – 44.447 ≈ –78.174
- Upper bound: –33.727 + 44.447 ≈ 10.720
- 95% CI: approximately (–78.17, 10.72)

8. Answer: No significant difference; 0 is contained in the interval (–78.17, 10.72).

- Check whether 0 (no difference) falls inside the confidence interval.
- Since 0 is between –78.17 and 10.72, we cannot reject the claim that the population mean difference equals 0.
- Conclusion: At the 5% significance level there is not enough evidence to conclude a significant difference in yields.

9. Answer: Mean diff ≈ 2.050; s_d ≈ 1.907; 95% CI ≈ (0.455, 3.645)

- Sum of differences: 3.1 + (–0.5) + 4.2 + 2.8 + 1.6 + 3.9 + (–1.2) + 2.5 = 16.4; mean = 16.4 / 8 = 2.05
- Compute each squared deviation from 2.05 and sum them: ≈ 25.38; s_d = sqrt(25.38 / 7) ≈ sqrt(3.626) ≈ 1.904
- Margin of error: 2.365 × 1.904 / sqrt(8) = 2.365 × 1.904 / 2.828 ≈ 2.365 × 0.673 ≈ 1.592
- 95% CI: (2.05 – 1.59, 2.05 + 1.59) ≈ (0.46, 3.64); since 0 is NOT in the interval, the dominant hand is significantly stronger.

10. Answer: 99% CI ≈ (–1.10, 17.70); evidence is inconclusive at 1% significance level.

- Margin of error: t* × s_d / sqrt(n) = 3.106 × 11.6 / sqrt(12) = 3.106 × 11.6 / 3.464 ≈ 3.106 × 3.350 ≈ 10.405
- Lower bound: 8.3 – 10.405 ≈ –2.105; Upper bound: 8.3 + 10.405 ≈ 18.705
- 99% CI ≈ (–2.11, 18.71)
- Since 0 is inside the interval, at the 1% significance level we cannot conclude the technique significantly increases scores. The researcher's claim is not supported at 99% confidence.

Scan to watch