Hypothesis Tests for the Means of Two Populations for dependent Samples

We used the t test to compare two sample means when the samples were independent. In this section, a different version of the t test is explained. This version is used when the samples are dependent. Samples are considered to be dependent samples when the subjects are paired or matched in some way. For example, suppose a medical researcher wants to see whether a drug will affect the reaction time of its users. To test this hypothesis, the researcher must pretest the subjects in the sample first. That is, they are given a test to ascertain their normal reaction times. Then after taking the drug, the subjects are tested again, using a posttest. Finally, the means of the two tests are compared to see whether there is a difference. Since the same subjects are used in both cases, the samples are related; subjects scoring high on the pretest will generally score high on the posttest, even after consuming the drug. Likewise, those scoring lower on the pretest will tend to score lower on the posttest. To take this effect into account, the researcher employs a t test, using the differences between the pretest values and the posttest values. Thus only the gain or loss in values is compared.

When the samples are dependent, a special t test for dependent means is used. This test employs the difference in values of the matched pairs. The hypotheses are as follows:

Two-tailed	Left-tailed	Right-tailed
$H_0: \mu_D = 0$ $H_1: \mu_D \neq 0$	$H_0: \mu_D = 0$ $H_1: \mu_D < 0$	$H_0: \mu_D = 0$ $H_1: \mu_D > 0$

Critical-Value and P-Value Approach

Formulas for the *t* Test for Dependent Samples

$$t = \frac{\overline{D} - \mu_D}{s_D/\sqrt{n}}$$

with d.f. $= n - 1$ and where

$$\overline{D} = \frac{\Sigma D}{n} \qquad \text{and} \qquad s_D = \sqrt{\frac{n\Sigma D^2 - (\Sigma D)^2}{n(n-1)}}$$

Confidence Interval Approach

Confidence Interval for the Mean Difference

$$\overline{D} - t_{\alpha/2}\frac{s_D}{\sqrt{n}} < \mu_D < \overline{D} + t_{\alpha/2}\frac{s_D}{\sqrt{n}}$$

d.f. $= n - 1$

Procedure Table

Testing the Difference Between Means for Dependent Samples

Step 1 State the hypotheses and identify the claim.

Step 2 Find the critical value(s).

Step 3 Compute the test value.

a. Make a table, as shown.

		A	B
X_1	X_2	$D = X_1 - X_2$	$D^2 = (X_1 - X_2)^2$
⋮	⋮		
		$\Sigma D =$ ______	$\Sigma D^2 =$ ______

b. Find the differences and place the results in column A.

$$D = X_1 - X_2$$

c. Find the mean of the differences.

$$\overline{D} = \frac{\Sigma D}{n}$$

d. Square the differences and place the results in column B. Complete the table.

$$D^2 = (X_1 - X_2)^2$$

e. Find the standard deviation of the differences.

$$s_D = \sqrt{\frac{n\Sigma D^2 - (\Sigma D)^2}{n(n-1)}}$$

f. Find the test value.

$$t = \frac{\overline{D} - \mu_D}{s_D/\sqrt{n}} \qquad \text{with d.f.} = n - 1$$

Step 4 Make the decision.

Step 5 Summarize the results.

Example:

A physical education director claims by taking a special vitamin, a weight lifter can increase his strength. Eight athletes are selected and given a test of strength, using the standard bench press. After 2 weeks of regular training, supplemented with the vitamin, they are tested again. Test the effectiveness of the vitamin regimen at ∝=0.05. Each value in these data represents the maximum number of pounds the athlete can bench-press. Assume that the variable is approximately normally distributed.

Athlete	1	2	3	4	5	6	7	8
Before (X_1)	210	230	182	205	262	253	219	216
After (X_2)	219	236	179	204	270	250	222	216

Solution

Step 1 State the hypotheses and identify the claim. For the vitamin to be effective, the before weights must be significantly less than the after weights; hence, the mean of the differences must be less than zero.

H_0: $\mu_D = 0$ and H_1: $\mu_D < 0$ (claim)

Step 2 Find the critical value. The degrees of freedom are $n - 1$. In this case, d.f. $= 8 - 1 = 7$. The critical value for a left-tailed test with $\alpha = 0.05$ is -1.895.

Step 3 Compute the test value.

a. Make a table.

		A	B
Before (X_1)	**After (X_2)**	$D = X_1 - X_2$	$D^2 = (X_1 - X_2)^2$
210	219	−9	81
230	236	−6	36
182	179	+3	9
205	204	+1	1
262	270	−8	64
253	250	+3	9
219	222	−3	9
216	216	0	0
		$\Sigma D = -19$	$\Sigma D^2 = 209$

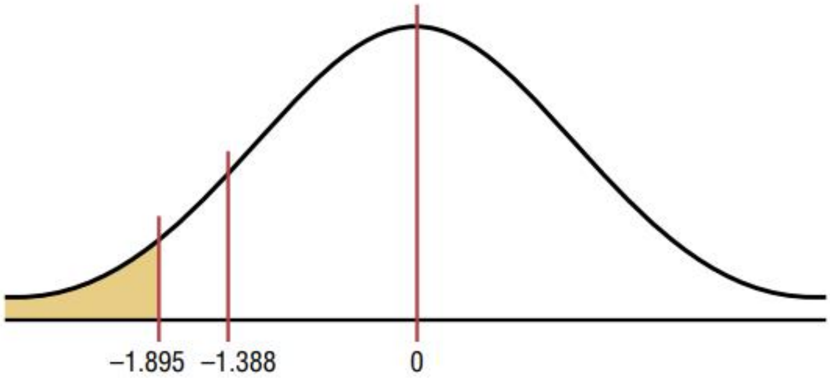
Find the standard deviation of the differences.

$$s_D = \sqrt{\frac{n\Sigma D^2 - (\Sigma D)^2}{n(n-1)}}$$
$$= \sqrt{\frac{8 \cdot 209 - (-19)^2}{8(8-1)}}$$
$$= \sqrt{\frac{1672 - 361}{56}}$$
$$= 4.84$$

Find the test value.

$$t = \frac{\overline{D} - \mu_D}{s_D/\sqrt{n}} = \frac{-2.375 - 0}{4.84/\sqrt{8}} = -1.388$$

Step 4 Make the decision. The decision is not to reject the null hypothesis at $\alpha = 0.05$, since $-1.388 > -1.895$, as shown in Figure 9–6.



Step 5 Summarize the results. There is not enough evidence to support the claim that the vitamin increases the strength of weight lifters.

Example:

A dietitian wishes to see if a person's cholesterol level will change if the diet is supplemented by a certain mineral. Six subjects were pretested, and then they took the mineral supplement for a 6-week period. The results are shown in the table. (Cholesterol level is measured in milligrams per deciliter.) Can it be concluded that the cholesterol level has been changed at ∝=0.10? Assume the variable is approximately normally distributed.

Subject	1	2	3	4	5	6
Before (X_1)	210	235	208	190	172	244
After (X_2)	190	170	210	188	173	228

Step 1 State the hypotheses and identify the claim. If the diet is effective, the before cholesterol levels should be different from the after levels.

H_0: $\mu_D = 0$ and H_1: $\mu_D \neq 0$ (claim)

Step 2 Find the critical value. The degrees of freedom are 5. At $\alpha = 0.10$, the critical values are ± 2.015.

Step 3 Compute the test value.

Before (X_1)	**After (X_2)**	**A** $D = X_1 - X_2$	**B** $D^2 = (X_1 - X_2)^2$
210	190	20	400
235	170	65	4225
208	210	−2	4
190	188	2	4
172	173	−1	1
244	228	16	256
		$\Sigma D = 100$	$\Sigma D^2 = 4890$

Find the mean of the differences.

$$\overline{D} = \frac{\Sigma D}{n} = \frac{100}{6} = 16.7$$

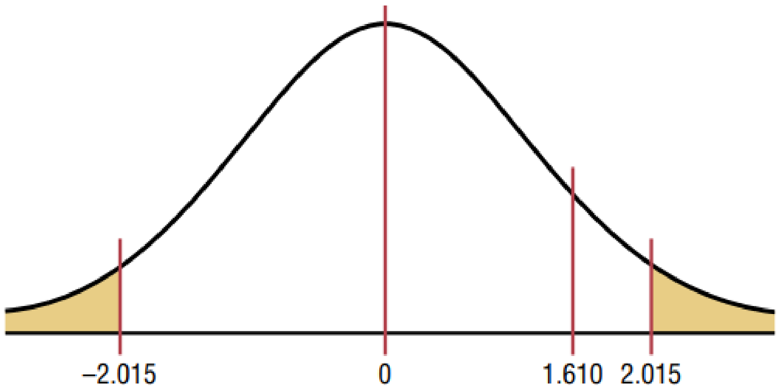
Find the standard deviation of the differences.

$$s_D = \sqrt{\frac{n\Sigma D^2 - (\Sigma D)^2}{n(n-1)}}$$
$$= \sqrt{\frac{6 \cdot 4890 - 100^2}{6(6-1)}}$$
$$= \sqrt{\frac{29{,}340 - 10{,}000}{30}}$$
$$= 25.4$$

Find the test value.

$$t = \frac{\overline{D} - \mu_D}{s_D/\sqrt{n}} = \frac{16.7 - 0}{25.4/\sqrt{6}} = 1.610$$

Step 4 Make the decision. The decision is to not reject the null hypothesis, since the test value 1.610 is in the noncritical region, as shown in Figure 9–7.



Step 5 Summarize the results. There is not enough evidence to support the claim that the mineral changes a person's cholesterol level.

Confidence Interval Approach

Find the 90% confidence interval for the data

Find the 90% confidence interval for the data in Example 9–7.

Solution

Substitute in the formula.

$$\overline{D} - t_{\alpha/2}\frac{s_D}{\sqrt{n}} < \mu_D < \overline{D} + t_{\alpha/2}\frac{s_D}{\sqrt{n}}$$

$$16.7 - 2.015 \cdot \frac{25.4}{\sqrt{6}} < \mu_D < 16.7 + 2.015 \cdot \frac{25.4}{\sqrt{6}}$$

$$16.7 - 20.89 < \mu_D < 16.7 + 20.89$$

$$-4.19 < \mu_D < 37.59$$

Since 0 is contained in the interval, the decision is to not reject the null hypothesis H_0: $\mu_D = 0$.