Applying the Concepts 9–3

Air Quality

As a researcher for the EPA, you have been asked to determine if the air quality in the United States has changed over the past 2 years. You select a random sample of 10 metropolitan areas and find the number of days each year that the areas failed to meet acceptable air quality standards. The data are shown.

Year 1	18	125	9	22	138	29	1	19	17	31
Year 2	24	152	13	21	152	23	6	31	34	20

Source: *The World Almanac and Book of Facts.*

Based on the data, answer the following questions.

1. What is the purpose of the study?
2. Are the samples independent or dependent?
3. What hypotheses would you use?
4. What is (are) the critical value(s) that you would use?
5. What statistical test would you use?
6. How many degrees of freedom are there?
7. What is your conclusion?
8. Could an independent means test have been used?
9. Do you think this was a good way to answer the original question?

See page 531 for the answers.

Exercises 9–3

1. Classify each as independent or dependent samples.

a. Heights of identical twins Dependent

b. Test scores of the same students in English and psychology Dependent

c. The effectiveness of two different brands of aspirin Independent

d. Effects of a drug on reaction time, measured by a before-and-after test Dependent

e. The effectiveness of two different diets on two different groups of individuals Independent

For Exercises 2 through 10, perform each of these steps. Assume that all variables are normally or approximately normally distributed.

a. State the hypotheses and identify the claim.

b. Find the critical value(s).

c. Compute the test value.

d. Make the decision.

e. Summarize the results.

Use the traditional method of hypothesis testing unless otherwise specified.

2. Retention Test Scores A sample of non-English majors at a selected college was used in a study to see if the student retained more from reading a 19th-century novel or by watching it in DVD form. Each student was assigned one novel to read and a different one to watch, and then they were given a 20-point written quiz on each novel. The test results are shown below. At $\alpha = 0.05$, can it be concluded that the book scores are higher than the DVD scores?

Book	90	80	90	75	80	90	84
DVD	85	72	80	80	70	75	80

3. Improving Study Habits As an aid for improving students' study habits, nine students were randomly selected to attend a seminar on the importance of education in life. The table shows the number of hours each student studied per week before and after the

Technology *Step by Step*

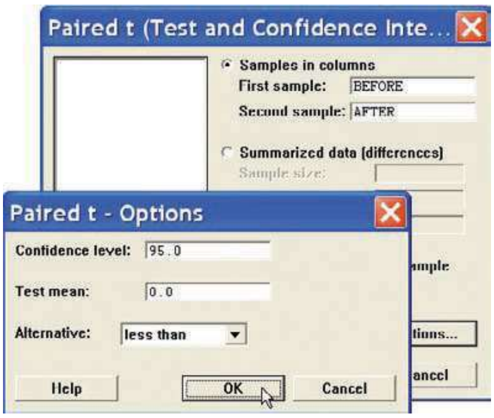
MINITAB
Step by Step

Test the Difference Between Two Means: Dependent Samples

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 $\alpha = 0.05$. 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

1. Enter the data into C1 and C2. Name the columns **Before** and **After.**
2. Select **Stat>Basic Statistics>Paired t.**
3. Double-click C1 Before for First sample.
4. Double-click C2 After for Second sample. The second sample will be subtracted from the first. The differences are not stored or displayed.
5. Click [Options].
6. Change the Alternative to less than.
7. Click [OK] twice.



Paired t-Test and CI: BEFORE, AFTER

```
Paired t for BEFORE - AFTER
             N      Mean     StDev    SE Mean
BEFORE       8   222.125    25.920      9.164
AFTER        8   224.500    27.908      9.867
Difference   8   -2.37500   4.83846     1.71065

95% upper bound for mean difference: 0.86597
t-Test of mean difference = 0 (vs < 0) : t-Value = -1.39  P-Value = 0.104.
```

Since the *P*-value is 0.104, do not reject the null hypothesis. The sample difference of −2.38 in the strength measurement is not statistically significant.

TI-83 Plus or TI-84 Plus
Step by Step

Hypothesis Test for the Difference Between Two Means: Dependent Samples

1. Enter the data values into L_1 and L_2.
2. Move the cursor to the top of the L_3 column so that L_3 is highlighted.
3. Type $\mathbf{L_1 - L_2}$, then press **ENTER.**
4. Press **STAT** and move the cursor to TESTS.
5. Press **2** for TTest.
6. Move the cursor to Data and press **ENTER.**

t-Test: Paired Two Sample for Means		
	Variable 1	*Variable 2*
Mean	31.875	29.625
Variance	6.696428571	14.55357143
Observations	8	8
Pearson Correlation	-0.757913399	
Hypothesized Mean Difference	0	
df	7	
t Stat	1.057517468	
P(T<=t) one-tail	0.1626994	
t Critical one-tail	1.894578604	
P(T<=t) two-tail	0.3253988	
t Critical two-tail	2.364624251	

Note: You may need to increase the column width to see all the results. To do this:

1. Highlight the columns D, E, and F.
2. Select **Format>AutoFit** Column Width.

The output shows a *P*-value of 0.3253988 for the two-tailed case. This value is greater than the alpha level of 0.05, so we fail to reject the null hypothesis.

9–4

Objective 4

Test the difference between two proportions.

Testing the Difference Between Proportions

The z test with some modifications can be used to test the equality of two proportions. For example, a researcher might ask, Is the proportion of men who exercise regularly less than the proportion of women who exercise regularly? Is there a difference in the percentage of students who own a personal computer and the percentage of nonstudents who own one? Is there a difference in the proportion of college graduates who pay cash for purchases and the proportion of non-college graduates who pay cash?

Recall from Chapter 7 that the symbol $\hat{p}$ ("p hat") is the sample proportion used to estimate the population proportion, denoted by p. For example, if in a sample of 30 college students, 9 are on probation, then the sample proportion is $\hat{p} = \frac{9}{30}$, or 0.3. The population proportion p is the number of all students who are on probation, divided by the number of students who attend the college. The formula for $\hat{p}$ is

$$\hat{p} = \frac{X}{n}$$

where

X = number of units that possess the characteristic of interest
n = sample size

When you are testing the difference between two population proportions p_1 and p_2, the hypotheses can be stated thus, if no difference between the proportions is hypothesized.

$$H_0: p_1 = p_2 \qquad \text{or} \qquad H_0: p_1 - p_2 = 0$$
$$H_1: p_1 \neq p_2 \qquad \qquad H_1: p_1 - p_2 \neq 0$$

Similar statements using < or > in the alternate hypothesis can be formed for one-tailed tests.

For two proportions, $\hat{p}_1 = X_1/n_1$ is used to estimate p_1 and $\hat{p}_2 = X_2/n_2$ is used to estimate p_2. The standard error of the difference is

$$\sigma_{\hat{p}_1-\hat{p}_2} = \sqrt{\sigma^2_{\hat{p}_1} + \sigma^2_{\hat{p}_2}} = \sqrt{\frac{p_1q_1}{n_1} + \frac{p_2q_2}{n_2}}$$