5. Inference for Population Means

5.1 Inference for the mean of a population using the t distribution

<u>Confidence interval for a population mean</u>

A one-sample confidence interval for a population mean μ is given by

$$\bar{x} \pm t^*_{n-1} \frac{s}{\sqrt{n}}$$

where t^*_{n-1} is the appropriate critical value from the t-distribution with n-1 degrees of freedom (see Topic 20, page 418 of the text). This procedure assumes the sample is a random sample from the population.

Example 5-1: Potato Chips. A potato chip manufacturing plant takes a random sample of 10 bags of chips from all the bags produced during one day and carefully measures the net weight of chips in each bag. The results are (in ounces):

11.91, 12.03, 11.98, 11.92, 11.96, 12.01, 11.93, 11.95, 12.06, 12.00

These data were entered using the SPSS Data Editor into a variable called ***weight***. To use these data to obtain a 99% confidence interval for the mean weight of all bags of chips produced at the plant that day, follow these steps.

1. Click **Analyze**, click **Descriptive Statistics**, and then click **Explore**. The SPSS window in Figure 5-1 appears.
2. Click ***weight***, then click  to move ***weight*** into the "Dependent List" box.
3. By default, a 95% confidence interval for μ will be computed. To change the confidence level, click **Statistics**. The "Explore Statistics" window shown in Figure 5-2 appears. Change 95 to **99** in the "Confidence Interval for Mean" box. Click **Continue**.
4. By default, the "Display" box in the lower left hand corner has "Both" selected. Click **Statistics**.
5. Click **OK**.

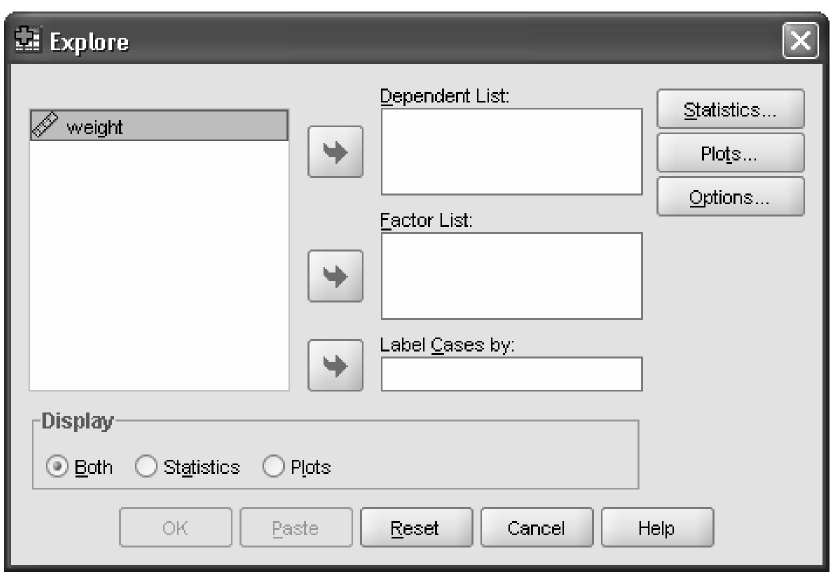


Figure 5-1

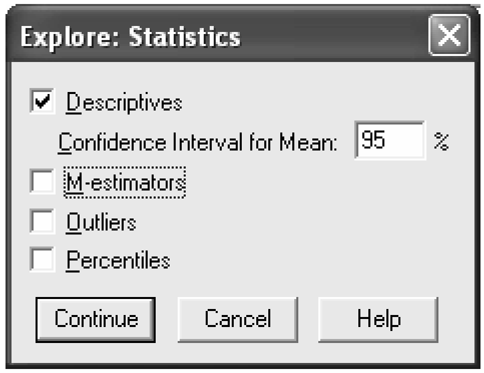


Figure 5-2

The resulting output is displayed in Table 5-1. A 99% confidence interval for the mean weight of all bags of chips produced that day is (11.92, 12.03) ounces. As stated on page 375 of the text, this confidence interval is only valid if the bags in the sample are chosen randomly from all the bags produced that day. The confidence interval formula also assumes that either the population is normally distributed or that the sample size is large ($n \geq 30$ or so). Unfortunately, with only 10 data values, it is difficult to judge normality. However, it is more important that the distribution not be too skewed and that there not be any outliers. That appears to be the case here (judging from a histogram, not shown) so we are OK. Past information can also be helpful in a case like this. The manufacturer has likely weighed many potato chip bags in the past and can use these past data to judge the appropriateness of the normality assumption.

Descriptives

			Statistic	Std. Error
weight	Mean		11.9750	.01572
	95% Confidence Interval for Mean	Lower Bound	11.9394	
		Upper Bound	12.0106	
	5% Trimmed Mean		11.9739	
	Median		11.9700	
	Variance		.002	
	Std. Deviation		.04972	
	Minimum		11.91	
	Maximum		12.06	
	Range		.15	
	Interquartile Range		.09	
	Skewness		.325	.687
	Kurtosis		-.954	1.334

Table 5-1

t test for a population mean

A significance test about a population mean μ can be carried out using a t test (see Topic 20, page 393 of the text). The test statistic is

$$t = \frac{\bar{x} - \mu_0}{s/\sqrt{n}}$$

where μ_0 represents the hypothesized value of the population mean. The t-distribution with n-1 degrees of freedom is used to compute the p-value.

Example 5-1: Potato Chips (page 73), continued. In the potato chip example, the weights of 10 randomly selected bags of chips from one day's production at a manufacturing plant were recorded. The net weight claimed on the packages from this plant is 12 ounces. Is there evidence that the mean net weight of bags produced that day is less than 12 ounces?

We can set this up as a hypothesis test about a mean. The hypotheses are

$$H_0: \mu = 12$$
$$H_a: \mu < 12$$

where μ is the mean weight of all bags of chips produced that day. To conduct the *t*-test in SPSS, follow these steps.

1. Click **Analyze**, then click **Compare Means**, and then click **One-Sample T Test**. The SPSS window in Figure 5-3 appears.
2. Click ***weight*** then click [arrow button] to move ***weight*** to the "Test Variable(s)" box.
3. Change 0 in the "Test Value" box to **12** (the value of μ under H_0).
4. Click **OK**.

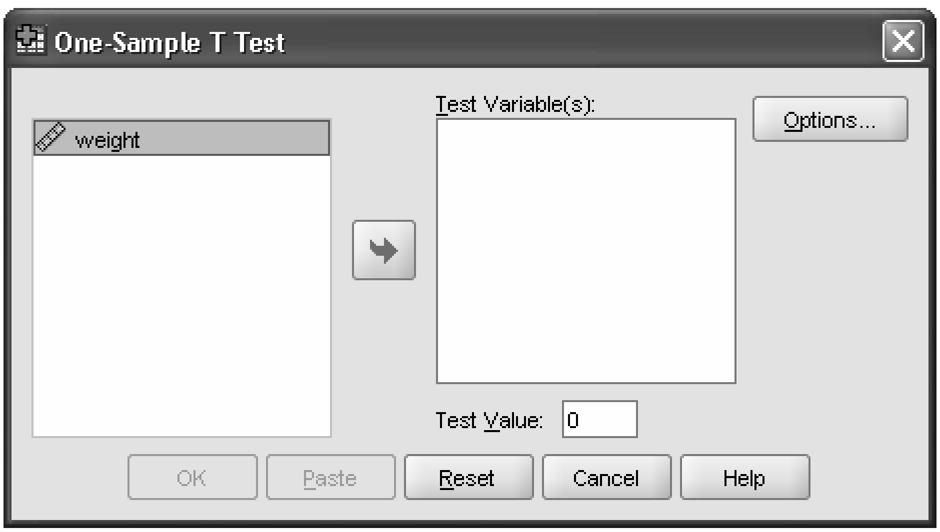


Figure 5-3

Table 5-2 and Table 5-3 show the output. From Table 5-2, the mean weight is 11.975 ounces with a standard deviation of .0497. From Table 5-3, the value of the *t* statistic is -1.590 and the p-value (called "Sig." for "Significance" in the SPSS output) for a two-sided test is .146. However, we wanted a one-sided test; in that case, the p-value is the area to the left of -1.590 for a *t* distribution with 9 degrees of freedom. Thus, the p-value is half of .146 or .073. Our conclusion would be that we have some evidence that the mean weight of all the bags is less than 12 ounces, but the evidence is not strong (see page 416 of the text for interpretation of p-values). See the discussion in the previous section, **Confidence interval for a population mean** (page 73), for the assumptions involved in a test of hypotheses.

Note: Table 5-3 also shows a confidence interval. This is not a confidence interval for the population mean μ, but is a confidence interval for the difference between μ and the hypothesized value of μ, that is, it's a confidence interval for $\mu - \mu_0$. You would need to add the value of μ_0 (12 in this case) to both endpoints of this confidence interval to get a confidence interval for μ.

One-Sample Statistics

	N	Mean	Std. Deviation	Std. Error Mean
weight	10	11.9750	.04972	.01572

Table 5-2

One-Sample Test

	Test Value = 12					
					95% Confidence Interval of the Difference	
	t	df	Sig. (2-tailed)	Mean Difference	Lower	Upper
weight	-1.590	9	.146	-.02500	-.0606	.0106

Table 5-3

Matched pairs t procedure

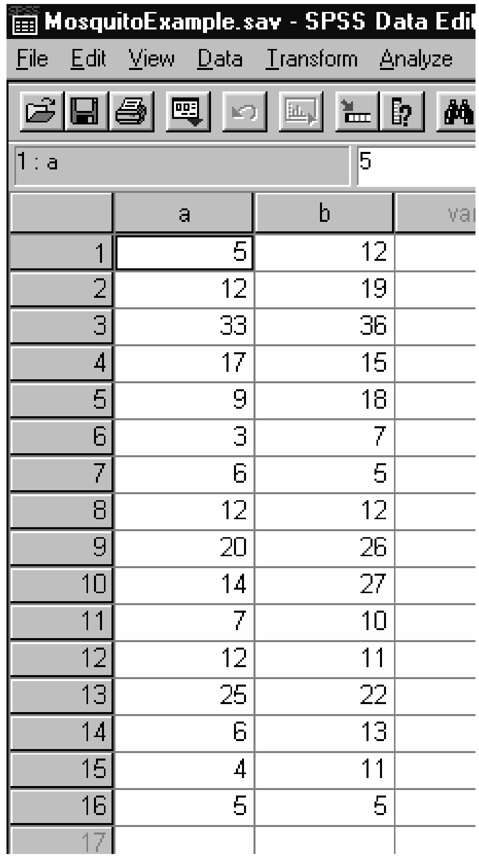
The one-sample *t* test can also be used for matched pairs data (also called "paired samples" data). The example below demonstrates this. See also Activity 23-5 on page 471 of the text.

Example 5-2: Mosquito repellents. Researchers were interested in comparing the effectiveness of two mosquito repellents. They had a group of 16 volunteers for the experiment. They considered two experimental designs. The first design would be to randomly divide the 16 subjects into two equal groups, apply repellent A to one group's arms and repellent B to the other group's arms. Then expose all subjects to a swarm of mosquitoes for a fixed period of time and compare the average numbers of bites for repellent A subjects to the average number for repellent B subjects. The second design would be to apply both repellents to each subject, one repellent on one arm (randomly choose which arm for each subject) and the other repellent on the other arm. Then compare the number of bites on the two arms. The experimenters chose the second design because it was known that people vary widely in their attractiveness to mosquitoes. Therefore, they felt it would be easier to detect a difference between the repellents if a matched pairs design was used. The results (number of bites on each arm) were:

Subject	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**	**13**	**14**	**15**	**16**
Repellent A	5	12	33	17	9	3	6	12	20	14	7	12	25	6	4	5
Repellent B	12	19	36	15	18	7	5	12	26	27	10	11	22	13	11	5

The goal is to carry out a test of the hypothesis that the population mean difference is 0 versus the alternative hypothesis that it is not 0 and to compute a 95% confidence interval for the population mean difference.

The data are entered with one subject per row and two variables, one representing the number of bites with repellent A and the other the number of bites with repellent B. The data as entered into SPSS are shown in Figure 5-4.



	a	b	var
1	5	12	
2	12	19	
3	33	36	
4	17	15	
5	9	18	
6	3	7	
7	6	5	
8	12	12	
9	20	26	
10	14	27	
11	7	10	
12	12	11	
13	25	22	
14	6	13	
15	4	11	
16	5	5	
17			

Figure 5-4

The analysis for the matched pairs experiment proceeds by computing the difference between the numbers of bites for each subject and then analyzing the differences. For example, we can compute a confidence interval for the mean population difference and carry out a one-sample *t* test to test the hypothesis that the population mean difference is 0. There are two ways to do this in SPSS. One is to create a new variable which is the difference between ***a*** and ***b*** (that is, ***a-b***) and then carry out a t test on the differences as described in the previous sections. The other way is to use the built-in SPSS Paired Samples T Test procedure that does the same thing but you do not need to have SPSS compute the difference ***a-b*** first. The advantage of computing the difference ***a-b*** is that then we can examine the distribution of the differences (with a histogram, for example). The assumption of the matched pairs t test is that the distribution of the *differences* (not of each variable by itself) is normal, so it is important that we examine the distribution of the differences.

To carry out the matched pairs t test the first way, follow these steps.

1. Create a new variable that is the difference between ***a*** and ***b*** (See Section 1.14, page 12): click **Transform** and click **Compute Variable**. Enter the name of the new variable (e.g., "diff") in the Target Variable box and enter the expression "a-b" (without the quotes) in the Numeric Expression box. Click **OK**.
2. Carry out a one-sample t test as described in the previous section (page 74): click **Analyze**, click **Compare Means** and click **One-Sample T Test**. Enter the variable ***diff*** into the "Test Variable(s)" box. The "Test Value" should be **0** (the default). Then click **OK**. The results are shown in Table 5-4 and Table 5-5.

One-Sample Statistics

	N	Mean	Std. Deviation	Std. Error Mean
DIFF	16	-3.6875	4.5712	1.1428

Table 5-4

One-Sample Test

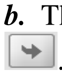
	Test Value = 0					
					95% Confidence Interval of the Difference	
	t	df	Sig. (2-tailed)	Mean Difference	Lower	Upper
DIFF	-3.227	15	.006	-3.6875	-6.1233	-1.2517

Table 5-5

We see that the mean observed difference (A minus B) in the number of bites was -3.69. This means that the repellent A arm had 3.69 *fewer* bites, on average, than the repellent B arm. A 95% confidence interval for the population mean difference is (-6.12, -1.25). A two-sided test of the hypothesis that the population mean difference is 0 gives a p-value of .006. This is strong evidence that the repellants differ and that repellent A is more effective, on average, than repellent B.

Since the "Test Value" is 0, the confidence interval reported in Table 5-5 is a confidence interval for the mean difference in number of bites for the two repellents. We are 95% confident that repellent A would lead to between 1.25 and 6.12 fewer bites on average on an arm in the fixed time period used in the test. A histogram of the variable ***diff*** indicates that the distribution is not severely skewed and there are no outliers so the t procedures are valid (see pages 395 & 465 of the text).

To carry out the matched pairs t test the second way, follow these steps. Although you do not have to compute the differences using this method, you should compute them anyway so that you can look at a histogram of the differences to see if the assumptions of the t-test are satisfied.

1. Click **Analyze**, click **Compare Means**, and click **Paired Samples T Test**. The window in Figure 5-5 appears.
2. Click ***a*** in the box on the left side. The variable ***a*** appears after "Variable 1" in the "Current Selections" box.
3. Click ***b.*** The variable ***b*** appears after "Variable 2" in the "Current Selections" box.
4. Click . Then ***a-b*** Appears in the "Paired Variables" box.
5. By default, a 95% confidence interval for the population mean difference will be part of the output. To change the confidence level, click **Options**, and change 95 to the desired confidence level.
6. Click **OK**.

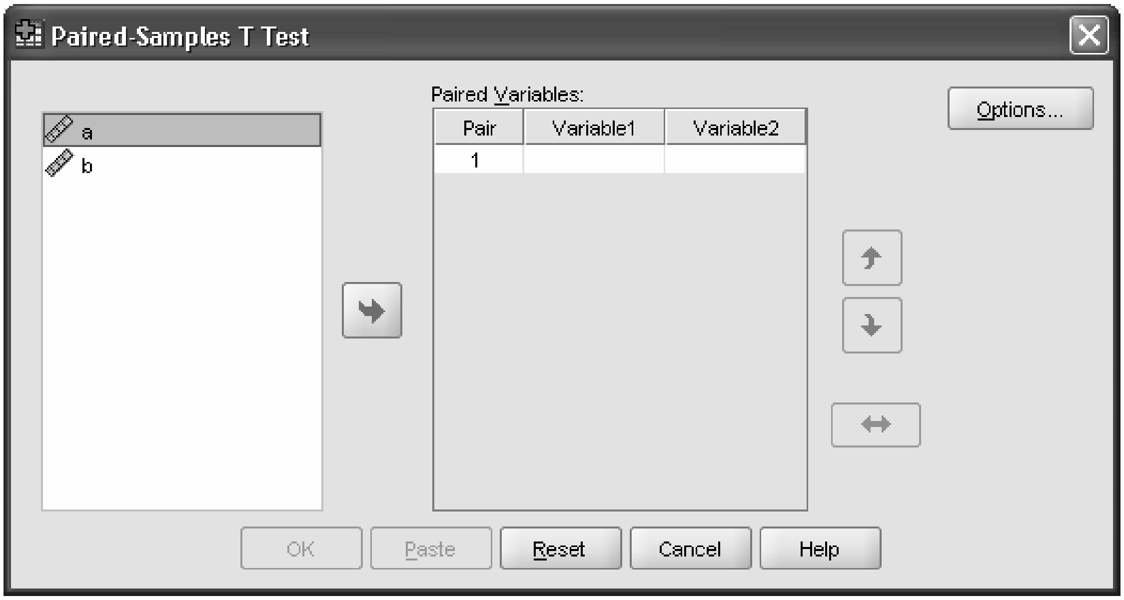


Figure 5-5

Table 5-6 is part of the resulting output. We get the same results as in Table 5-4 and Table 5-5.

Paired Samples Test

	Paired Differences					t	df	Sig. (2-tailed)
				95% Confidence Interval of the Difference				
	Mean	Std. Deviation	Std. Error Mean	Lower	Upper			
Pair 1 A - B	-3.69	4.57	1.14	-6.12	-1.25	-3.227	15	.006

Table 5-6

5.2 Comparing population means: the two-sample t procedure

The two-sample *t* procedures are used for making inferences about the difference between two population means. They are used when data are collected from independent random samples from the two populations.

Example 5-3: Blood Pressure and Calcium Intake. Examinations of large numbers of people have revealed a relationship between calcium intake and blood pressure, with higher calcium intakes being associated with lower blood pressures. The relationship is particularly strong among black men. Since these data are from observational studies and not controlled experiments, one cannot conclude that higher calcium intake *causes* lower blood pressure. Therefore, a controlled experiment is designed to investigate this relationship in black men. The subjects are 21 healthy black men. A randomly chosen group of 10 of the men receive a calcium supplement for 12 weeks. The control group of 11 men receive a placebo pill that looked identical. The experiment is double-blind. The response variable is the decrease in systolic blood pressure for a subject after 12 weeks, in millimeters of mercury. A negative value indicates an increase in blood pressure. The results are:

Calcium group: 7, -4, 18, 17, -3, -5, 1, 10, 11, -2
Placebo group: -1, 12, -1, -3, 3, -5, 5, 2, -11, -1, -3

Calculate a 99% confidence interval for the difference between the population mean blood pressure changes and carry out a two-sided test of the hypothesis that there is no difference between the population means.

The data were entered in the SPSS Data Editor as 21 cases and two variables: ***group*** (calcium or placebo) and ***response***. Some of the data as they appear in the Data Editor are shown in Figure 5-6. The variable ***group*** has been coded as 1 for calcium and 2 for placebo. Value labels (Section 1.9) for ***group*** identify which group is which.

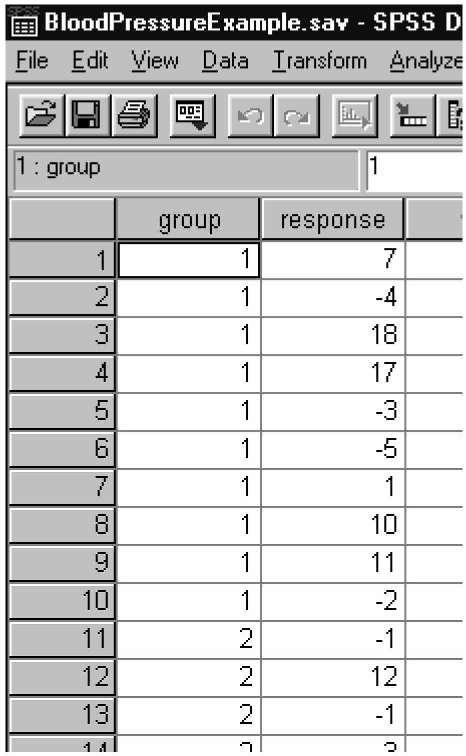


Figure 5-6

To calculate a confidence interval for the difference in population means and to carry out the hypothesis test, follow these steps (both will be done simultaneously).

1. Click **Analyze**, click **Compare Means**, and then click **Independent Samples T Test**. The SPSS window in Figure 5-7 appears.
2. Click ***response***, then click [→] to move ***response*** into the "Test Variable(s)" box.
3. Click ***group***, then click [→] to move ***group*** into the "Grouping Variable" box.
4. Click **Define Groups**. The SPSS window in Figure 5-8 appears.
5. Type **1** in the "Group 1" box. Press the tab key. Type **2** in the "Group 2" box. Then click **Continue**. Note: the groups must be defined by the values that appear in the SPSS Data Editor, not by the value labels. For example, if we had made group a string variable with values **C** and **P**, then we would enter **C** and **P** into these boxes.
6. By default, a 95% confidence interval for the difference in population means ($\mu_1 - \mu_2$) will be calculated. To change the confidence level to 99%, click **Options**, change 95 to **99** in the "Confidence Interval" box, and then click **Continue**.
7. Click **OK**.

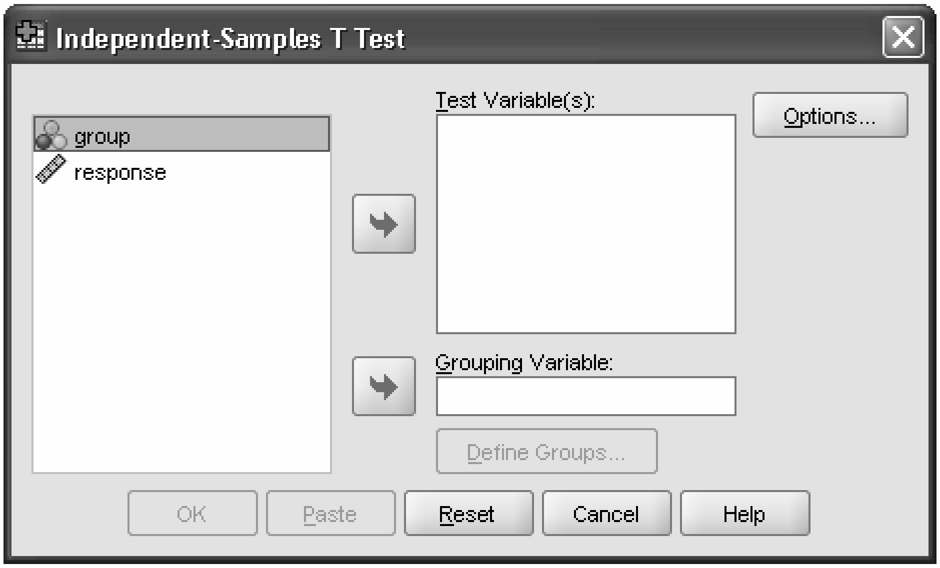


Figure 5-7

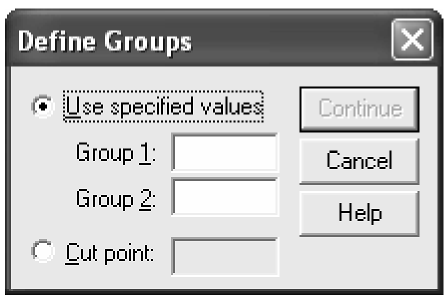


Figure 5-8

Table 5-7 and Table 5-8 are the resulting output. Table 5-8 gives the result of two different *t* procedures: one which assumes that the variances of the two populations are equal and one which does not. You should ignore the "Equal variances assumed" result and *always* use the "Equal variances not assumed" result. The "Equal variances not assumed" procedure is the one described in the text in Topic 25. You should also ignore "Levene's Test for Equality of Variances". The end result is that a 99% confidence interval for the difference (Calcium minus Placebo) in population mean blood pressure decrease is (-4.36, 14.91). This means that we are 99% confident that Calcium is anywhere from 4.36 units *worse* than the Placebo to 14.91 units *better* than the Placebo in terms of average systolic blood pressure decrease. Not surprisingly, then, the p-value for the test of equality of the population means is .129. Thus, there is little evidence of a difference between Calcium and the Placebo. (Note: this does not mean we have strong evidence that there is no difference, only that we lack evidence that there is a difference. A larger study could well find a significant difference.) Final note: the degrees of freedom SPSS uses for the two-sample t test with equal variances not assumed is different from what the text uses in Topic 25. The text uses the smaller of $n_1 - 1$ and $n_2 - 1$. The degrees of freedom used by SPSS is based on a somewhat complicated formula which depends on the samples sizes and the sample variances. Using the degrees of freedom from this formula gives a more accurate test and confidence interval and is preferred to the simpler formula in our text (which is used to avoid complicated hand calculations).

Group Statistics

	GROUP	N	Mean	Std. Deviation	Std. Error Mean
RESPONSE	Calcium	10	5.00	8.74	2.76
	Placebo	11	-.27	5.90	1.78

Table 5-7

Independent Samples Test

		Levene's Test for Equality of Variances		t-test for Equality of Means						
									99% Confidence Interval of the Difference	
		F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	Lower	Upper
RESPONSE	Equal variances assumed	4.351	.051	1.634	19	.119	5.27	3.23	-3.96	14.50
	Equal variances not assumed			1.604	15.591	.129	5.27	3.29	-4.36	14.91

Table 5-8