CHAPTER EIGHT

Two-Sample Means

8.0 Introduction to Two-Sample Hypothesis Testing

In previous sections, the one-sample comparison of the sample mean to some hypothesized population mean was discussed. Although the one-sample test for mean is occasionally explored in some research, most research studies involve the comparison of two or more samples. For example, an educator may which to study the performance of male and female on a particular aptitude test. In this study the researcher may be concern with the comparison of the means of each gender or the mean difference between them.

In this section we will examine the case where two samples means, M_1 and M_2 are used to make inference about two population means, μ_1 and μ_2. We will study three cases in this chapter: 1. two independent samples from different populations with homogeneous variances, 2. two independent samples from different populations with non-homogeneous variances, and 3. two samples that or correlated or related. Comparing means from two samples that or correlated is the same as dependent sample comparison analysis.

The **independent** samples comparison of means is often samples from two populations that may have the same variance or distribution. This is called the **homogeneous variances** case. When two variances from independent samples are not the same (or from different distribution), we call this situation the **non-homogeneous**

variances case. When the samples are related or correlated, the comparison of the sampling distributions of the means is called the dependent or correlated case.

The hypotheses statements for the two-sample case for the means involve statements about the difference of the means. The null hypothesis statement is that there is no difference between the population means and is

H_0: $\mu_1 - \mu_2 = 0$ (no difference between population means)

The population mean μ_2 may be a population mean estimated from a sample mean or a hypothesized value, we call the **test value**. The alternative hypothesis states that there is a difference between the two population means. The alternative hypothesis can be stated in one of three ways:

H_a: $\mu_1 - \mu_2 \neq 0$ or $\mu_1 \neq \mu_2$ (there is a difference between population means) ***or***

H_a: $\mu_1 > \mu_2$ (mean 1 is greater than mean two) ***or***

H_a: $\mu_1 < \mu_2$ (mean 1 is less than mean 2)

There are two approaches to finding the standard error based on contributions of both samples. One approach involves pooling the variances of both samples and the other approach involves combining the standard errors of both and adjusting the degree of freedom if the variances are non-homogeneous. The formula that combines the standard errors of the two variances is an estimated standard error of the mean difference:

standard error of mean difference = $s_d = \sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}$ (definition formula)

$s_d = \sqrt{\frac{n_1 s_1^2 + n_1 s_1^2}{n_1 + n_2 - 2}(\frac{1}{n_1} + \frac{1}{n_2})}$ (computational formula)

To decide on which approach to use when testing hypotheses involving two sample means, we follow the strategy outlined in Figure 8.0.1. Figure 8.0.1 shows the steps used to decide on whether to use the independent homogeneous variances or the independent non-homogeneous or the dependent *t* test strategy developed in this section and the next two sections.

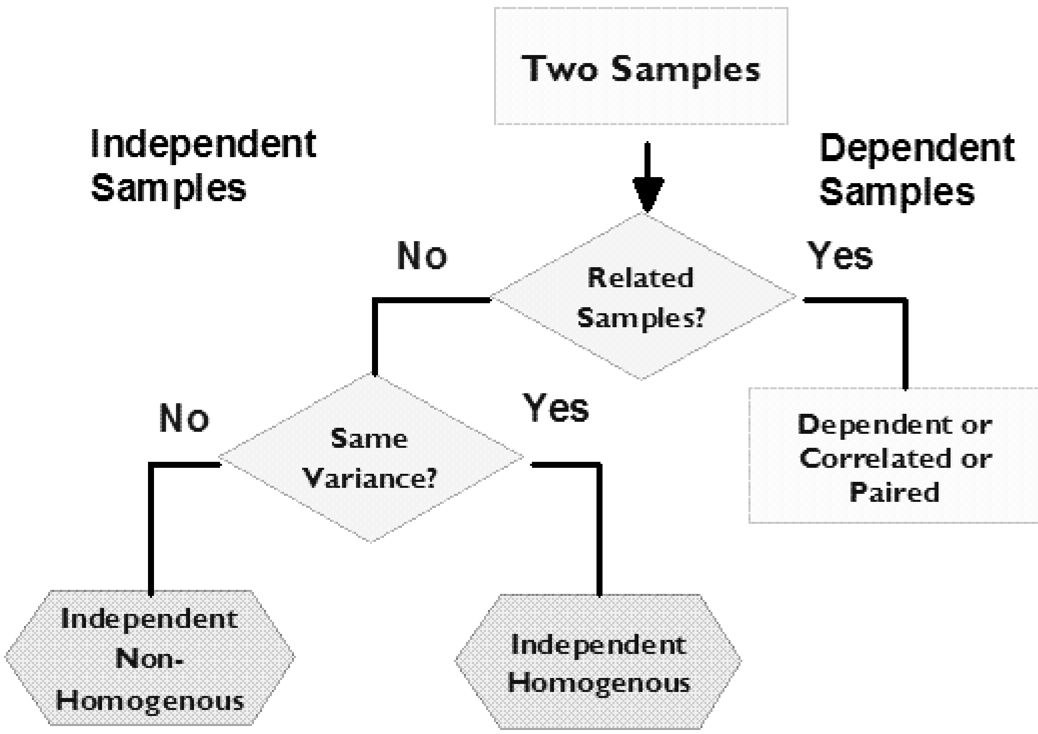

Figure 8.0.1 Decision Tree for Hypotheses Involving Two Samples

Table 8.0.1 summarizes the statistics and parameters used to evaluate both one sample and two samples hypothesis testing about means. There are several ways of stating the null hypothesis and test statistics for the two-sample means test. Notice that the more descriptive statement of the two-sample *t* statistics is simplified from the formula. Formulas for computing the standard error, s_{error} for each case when comparing two samples means will be introduced when cases are presented.

$$t = \frac{\text{mean difference}}{\text{standard error}} = \frac{(M_1 - M_2) - 0}{s_{error}}$$

Table 8.0.1 *Basic Statistics and Parameters for Comparing Means*

	Sample Statistics	Population Parameter	Estimated Standard Error	Test Statistics
One-Sample *t* test	M	μ	$\sqrt{\frac{s^2}{n}}$	$t = \frac{M - \mu}{s_M}$
Two-Sample *t* test	$(M_1 - M_2)$	μ_1 - μ_2	$\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}$	$t = \frac{M_1 - M_2}{s_{error}}$

Degrees of Freedom

The degrees of freedom (***df***) describes the number of scores in a sample that are allowed to vary. The sample size for sample one is denoted by the constant ***n***; for two samples, the first sample size is denoted by n_1 and for the second sample it is denoted by n_2. For the one-sample *t* test for comparing the mean to some hypothesized mean or test value, the degrees of freedom is $df = n - 1$. The degrees of freedom for the two-sample *t* test for comparing means from samples with homogeneous (same) variances is $df = n_1 + n_2 - 2$. The degrees of freedom for the two-sample correlated *t* test or paired samples is $df = n - 1$, where ***n*** is number of paired set of data scores for both samples. The degrees of freedom for the two-sample *t* test for comparing means from samples with non-homogeneous variances is given by the formula below. Table 8.0.2 summaries the formulas used to compute the degrees of freedom for the *t* test.

$$df = \frac{\left(\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}\right)^2}{\frac{\left(\frac{s_1^2}{n_1}\right)^2}{n_1 - 1} + \frac{\left(\frac{s_2^2}{n_2}\right)^2}{n_2 - 1}}$$

(Independent samples with non-homogeneous variances)

Table 8.0.2 *Degrees of Freedom Summary for t Test*

Types of *t* Test	**Degrees of Freedom Formula**
One Sample	$n - 1$
Two Independent Samples with Homogeneous Variances	$n_1 + n_2 - 2$
Two Independent Samples with Non-homogeneous Variances	$df = \frac{\left(\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}\right)^2}{\frac{\left(\frac{s_1^2}{n_1}\right)^2}{n_1 - 1} + \frac{\left(\frac{s_2^2}{n_2}\right)^2}{n_2 - 1}}$
Two Dependent Samples (Correlated Samples)	$n - 1$

Test for Homogeneity of Variances

Often it is necessary to compare two population variances to see if they are the same or homogeneous. The sample variance, s^2 is used to make inference about population variance, s^2. The underlying sampling distribution for comparing two population variances is the ***F* distribution**. The hypothesis test for comparing two population variances has the following hypotheses:

H_0: $s_1^{\ 2} = s_2^{\ 2}$ and the alternative hypothesis is either

H_a: $s_1^{\ 2} \neq s_2^{\ 2}$ ***or*** Ha: $s_1^{\ 2} = s_2^{\ 2}$ ***or*** H_a: $s_1^{\ 2} = s_2^{\ 2}$

Since all we need here is to know if the variances are the same or not, the alternative hypothesis stated as H_a: $s_1^{\ 2} \neq s_2^{\ 2}$ will be used in this chapter. The *F* statistics for hypothesis test about two population variances with $s_1^{\ 2} = s_2^{\ 2}$ based on sample variances, $s_1^{\ 2} = s_2^{\ 2}$ is

$$F = \frac{\text{sample one variance}}{\text{sample two variance}} = \frac{s_1^2}{s_2^2}$$

The ratio of the sample variances inputted in the *F* statistics formula is such that the larger variance is in the numerator and the smaller is in the denominator. There are many *F* distributions based on the degrees of freedom of the samples and the significance level, alpha. The critical value for the *F* distribution at alpha, $a = 0.05$ for two samples of sizes $n_1 = 16$ and $n_2 = 21$ is $\boldsymbol{F_{CV} = 2.20}$ ($df_1 = n_1 - 1 = 15$ and $df_2 = n_2 - 1 = 20$). The F_{CV} can be obtained from Table A6 and Table A7, alpha 0.05 and 0.01 respectively. Figure 8.0.2 shows an *F* distribution with $a = 0.05$, $df_1 = 15$ (degrees of freedom of the numerator) and $df_2 = 20$ (degrees of freedom of the denominator). The *F* statistics for two samples of $n_1 = 16$, $s_1^{\ 2} = 12$ and $n_2 = 21$, $s_2^{\ 2} = 5$ is

$$F \text{ statistics} = F = \frac{s_1^2}{s_2^2} = \frac{12}{5} = 2.4$$

Because $F_{statistics} = 2.4 > F_{CV} = 2.2$, we reject the null hypothesis that the two population variances are the same. The *p* value = 0.035 (From Excel programs) and since this is less than $a = 0.05$ we reject the null hypothesis that the variances are the same. Table 8.0.3 shows scores for two independent samples with equal variances. The next section evaluates the case for comparing two population means with same variances.

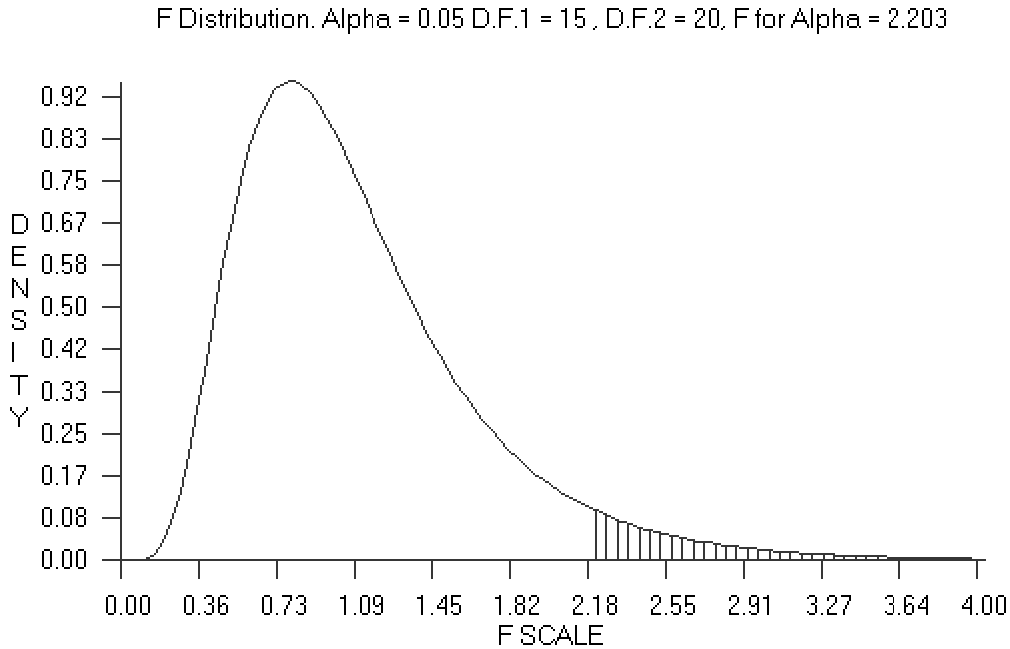

Figure 8.0.2 A *F* Distribution: $a = 0.05$, $df_1 = 15$, $df_2 = 20$

Table 8.0.3 *Two Independent Samples: Example 1*

Sample 1	Sample 2
85	79
73	73
68	45
65	81
62	71
40	42
72	71
68	79
63	72
59	55
56	73
77	76
74	71
74	66
66	64
37	48

8.1 Two Independent Samples with Homogeneous Variances

The next procedure evaluates the hypotheses H_0: $\mu_1 - \mu_2 = 0$ and H_a: $\mu_1 - \mu_2 \neq 0$ (non-directional). First the F test is used to test the hypothesis that the sample variances are equal. From the following procedure we conclude that the variances are the same and we don't reject the null hypothesis that the means are the same (mean difference equal zero).

Problem Statement: Test the hypothesis that the population means for the two independent samples in Table 8.0.3 are equal at the alpha = 0.05 level.

Test for Homogeneity of Variances

F statistics $= F = \frac{s_1^2}{s_2^2} = \frac{158.46}{155.45} = 1.02$; $F_{CV} = 2.4$ (α =0.05, $df_1 = df_2 = 15$)

Test concludes **equal variance**s since $F_{stat} = 1.02 < F_{CV} =$ **2.4** (Table A6)

Step 1: State the hypotheses

H_0: $\mu_1 - \mu_2 = 0$ and H_a: : $\mu_1 - \mu_2 \neq 0$ (non-directional)

Step 2: Select the significance level

$a = 0.05$ (0.95 confidence interval)

Step 3: Compute the test statistics

standard error = $s_d = \sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}} = \sqrt{\frac{158.46}{16} + \frac{155.45}{16}} = 4.43$

test statistics = $t = \frac{M_1 - M_2}{s_d} = \frac{158.46 - 155.45}{4.43} = 0.381$

Mean Difference = 158.46 – 155.45 = **1.6875**

$df = n_1 + n_2 - 2 = 16 + 16 - 2 =$ **30**

Step 4: Determine the criteria for rejecting H_0

The t critical value: t_{CV} = **2.0423** (Table A2: 2-tailed, df = 30, a = 0.05) ***or***

CI_{95} (Mean Diff) = 1.6875 ± 2.0423(4.43) = **-10.74** to **7.36** ***or***

The p value is **0.706** from Excel: TDIST(0.381, 30, 2)

Step 5: Make a decision

We don't reject the null hypothesis for any of the following reasons:

1. The $t_{stat} = 0.381 < t_{CV} = 2.042$ ***or***
2. The confidence interval for mean difference does contain **zero** ***or***
3. The p value for the test statistics is > than $a = 0.05$

Therefore, there is no significant reason to reject the null hypothesis that the population means are the same.

Table 8.1.1 shows scores for two samples with homogeneous variances. The procedure that follows evaluates the hypotheses H_0: $\mu_1 - \mu_2 = 0$ and H_a: : $\mu_1 - \mu_2 \neq 0$ (non-directional). From the procedure we conclude that the variances are the same and we reject the null hypothesis and conclude that the population means are different.

Table 8.1.1 *Two Independent Samples: Example 2*

Sample 1	Sample 2
124	108
145	133
127	109
123	118
111	94
115	111
129	107
143	125
135	120
139	119

Problem Statement: Test the hypothesis that the population means of the two independent samples in Table 8.1.1 are equal at the alpha = 0.05.

Test for Homogeneity of Variances

$$F \text{ statistics} = F = \frac{s_1^2}{s_2^2} = \frac{130.32}{119.6} = 1.09;\ F_{CV} = 3.18\ (\alpha = 0.05,\ df_1 = df_2 = 9)$$

Test conclude **equal variance**s since $F_{stat} = 1.09 < F_{CV} =$ **3.18**

Step 1: State the hypotheses

H_0: $\mu_1 - \mu_2 = 0$ and H_a: : $\mu_1 - \mu_2 \neq 0$ (non-directional)

Step 2: Select the significance level

$a = 0.05$ (0.95 confidence interval)

Step 3: Compute the test statistics

$$\text{standard error} = s_d = \sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}} = \sqrt{\frac{130.32}{10} + \frac{119.6}{10}} = 5.0$$

$$\text{test statistics} = t = \frac{M_1 - M_2}{s_d} = \frac{129.1 - 114.4}{5} = 2.94$$

Mean Difference = 129.1 – 114.4 = **14.7**

$df = n_1 + n_2 - 2 = 10 + 10 - 2 =$ **18**

Step 4: Determine the criteria for rejecting H_0

The t critical value: $t_{CV} =$ **2.1** (Table A2: 2-tailed, $df = 18$, $a = 0.05$) ***or***

CI_{95} (Mean Diff) = 14.7 ± 2.1(5) = **4.20** to **25.2** ***or***

The p value is **0.009** from Excel: TDIST(2.94, 18, 2)

Step 5: Make a decision

We **reject** the null hypothesis for any of the following reasons:

1. The $t_{stat} = 2.94 > t_{CV} = 2.1$
2. The confidence interval for mean difference does not contain zero ***or***
3. The p value for the test statistics is $< a = 0.05$

Therefore, the population means are different.

Effect Size and *r* Square

The **effect size** is a measurement that produces a standardized mean difference for our test: the mean difference divided by the standard deviation. The mean difference for two samples is simply $M_1 - M_2$, where M_1 is the mean for sample one and M_2 is the mean for sample two. The **pooled standard deviation** is used to estimate the pooled population standard deviation for the independent samples case. Cohen's *d* was defined earlier as

$$d = \frac{\text{mean difference}}{\text{standard deviation}} = \frac{M_1 - M_2}{s_p},\ \text{where } s_p \text{ is the pooled standard deviation}$$

The pooled standard deviation for two samples (X_1 and X_2) is the square root of the pooled variance. The samples pooled variance provides a best estimate of the populations' pooled variance. A computational formula for the samples pooled standard deviation is

$$\sigma_p^2 \simeq s_p^2\text{, so } s_{pooled} = \sqrt{\frac{(\Sigma X_1^2 - \frac{(\Sigma X_1)^2}{n_1}) + (\Sigma X_2^2 - \frac{(\Sigma X_2)^2}{n_2})}{n_1 + n_2 - 2}}$$

The spooled standard deviation from the two samples in Table 8.1.1 is **11.18**. The summations statistics for computing the pooled standard deviation is shown in Table 8.1.2. This calculation can be easily performed by an excel spreadsheet or a calculator.

$$s_{pooled} = \sqrt{\frac{(\Sigma X_1^2 - \frac{(\Sigma X_1)^2}{n_1}) + (\Sigma X_2^2 - \frac{(\Sigma X_2)^2}{n_2})}{n_1 + n_2 - 2}} = \sqrt{\frac{(167841 - \frac{1666681}{10}) + (131950 - \frac{1308736}{10})}{10 = 10 - 2}} = 11.1786$$

The effect size is then the mean difference, **14.7** (M_1 = 129.1 **–** M_2 = 114.4) divided by the spooled standard deviation, so *d* = **1.32**. This is a large effect.

$$\text{effect size} = d = \frac{M_1 - M_2}{s_p} = \frac{14.7}{11.18} = 1.32$$

Table 8.1.2 *Summations for Pooled Variance Statistics: Example 2*

Sample 1, X_1	Sample 2, X_2	X_1^2	X_2^2
124	108	15376	11664
145	133	21025	17689
127	109	16129	11881
123	118	15129	13924
111	94	12321	8836
115	111	13225	12321
129	107	16641	11449
143	125	20449	15625
135	120	18225	14400
139	119	19321	14161
$\sum X_1 = 1291$	$\sum X_2 = 1144$	$\sum X_1^2 = 67841$	$\sum X_2^2 = 131950$
$(\sum X_1)^2 = 1666681$	$(\sum X_2)^2 = 1308736$		

Computing the t statistics for the independent measure of samples means allow for computing the percentage of variance accounted for by r^2. This we will call ***r* square** and it is similar to the coefficient of determination. The r square measures how much of the variability in the scores can be explained by the treatment effects or the difference between the treatments between sample one and sample two. For Example 2 above, the r square is 0.32 as computed from the t statistic and degree of freedom, df for the two independent samples means. Therefore, 32% of the variability is accounted for its value gives a measure of how big the effect of treatment is between both samples. In a practical sense, r square is another way of measuring the effect size, d.

$$r^2 = \frac{t^2}{t^2 + df} = \frac{(2.9405)^2}{(2.9405)^2 + 18} = \frac{8.6465}{26.64654025} = 0.32$$

Figure 8.1.1 shows the SPSS outputs for the two independent samples analysis for the data in Table 8.1.1. Figure 8.1.2 shows the SPSS procedure for evaluating the two-sample *t* test for the two-sample means. The *t* statistics, in particular the degrees of freedom, to use when making inference about the two population means depends upon the results of the homogeneity of variances test (Levene's Test). The SPSS analysis outputs are similar to the results shown from our analysis above. SPSS analysis outputs provide two tables summaries. The *Group Statistics* shows the means, standard deviations, sample sizes, and the standard error of the mean for each sample. The *Independent Samples Test* table provides testing statistics for evaluating either the case when the variances are equal or not equal. Statistics are provided that allow for evaluating both the homogeneity of variances and the equality of means hypotheses.

Group Statistics

	Y	N	Mean	Std. Deviation	Std. Error Mean
X	2.00	10	129.1000	11.41588	3.61002
	1.00	10	114.4000	10.93618	3.45832

Independent Samples Test

		Levene's Test for Equality of Variances		t-test for Equality of Means					95% Confidence Interval of the Difference	
		F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	Lower	Upper
X	**Equal variances** assumed	.036	.852	**2.940**	18	**.009**	14.70000	4.99922	4.19702	25.20298
	Equal variances not assumed			2.940	17.967	.009	14.70000	4.99922	4.19564	25.20436

Figure 8.1.1 SPSS Independent Samples Test Output

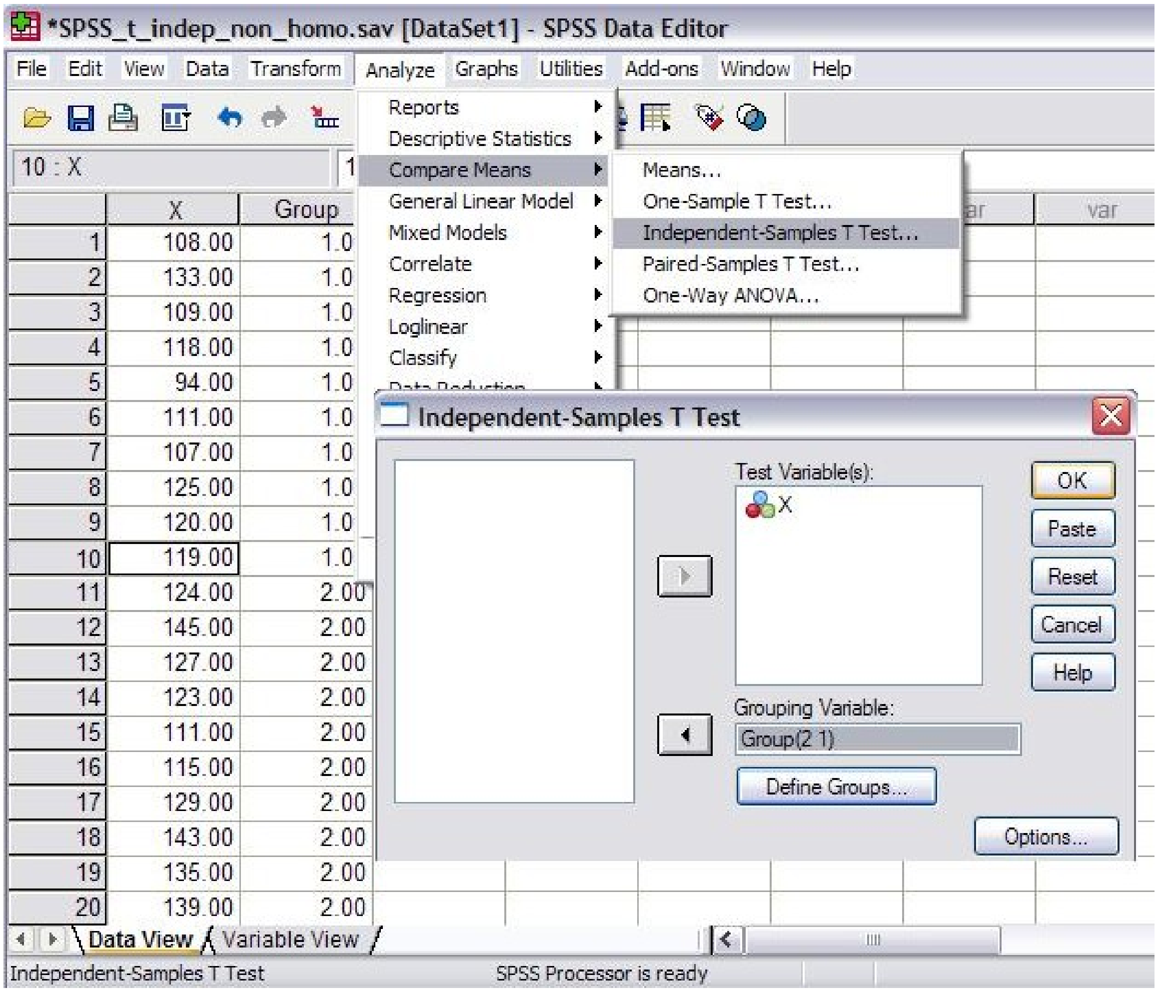

Figure 8.1.2 SPSS Procedure for Two-Sample Means Test

From the SPSS analysis of the data in Table 8.1.1, the null hypothesis is rejected with similar results as computed in the procedure summary above. There are separate excel statistical programs as part of this text that compute the *t* and the appropriate statistics for evaluating the null hypothesis for all three cases for comparing the means of two samples. There are also programs for computing the *F* statistics for homogeneity of variance test (equivalent to the SPSS Levene’s Test) and for computing the effect size and r square for measuring the percentage of variability accounted for by the treatment between sample one and sample two.