A **hypothesis test** is a statistical method that uses sample data to evaluate a hypothesis about a population or populations.

In very simple terms, the logic underlying the hypothesis-testing procedure is as follows:

1. **State the null and alternative hypotheses.** Usually the hypotheses concern the value of a population parameter such as the population mean.

2. **Set the criteria for a decision.** We use the null hypothesis to predict the characteristics that the sample should have.

3. **Collect data and compute sample statistics.**

4. **Make a decision.** We compare the obtained sample data with the prediction that was made from the null hypothesis. If the sample mean is consistent with the prediction, we conclude the null hypothesis is reasonable. But if there is a big discrepancy between the data and the prediction, we decide that the null hypothesis is wrong.

Hypotheses

There are two types of hypothesis statements. The two are complements of each other.

Null Hypothesis

- X is not a significant predictor of Y.
- X is not related to Y.
- X1 is not related to Y when controlling for X2 and X3.

Alternative Hypothesis

- X is a significant predictor of Y.
- X is related to Y.
- X1 is related to Y when controlling for X2 and X3.

Hypotheses Examples

Null or Alternative?	Statement
	Gender is not a significant predictor of income.
	Gender is a significant predictor of income when controlling for education.
	The exercise intervention will not change the BMI of healthy adults.

Errors in Hypothesis Testing

Type I Error – rejecting the null hypothesis when the null hypothesis is true.

Alpha – the probability of making a Type I error; the researcher sets the alpha level.

Type II Error – retaining the null hypothesis when the null hypothesis is false.

Beta – the probability of making a Type II error; the researcher can *reduce* the risk of a Type II error by increasing *power*.

Power – the probability of rejecting the null hypothesis when the null hypothesis is false. We want to have power levels of .80 or higher.

Beta	Power = 1 - Beta
.20	
.50	
	.80

Factors That Influence Power

Alpha level. Using a more liberal alpha level (e.g., .05 instead of .01) yields a slightly more powerful test. The drawback is that a more liberal alpha level has a higher risk of a Type I error.

Directional Hypothesis. A directional hypothesis yields slightly more power for a statistical test *if* the data match the predicted direction. *However,* if the predicted direction does not match the data then power approaches zero.

Sample Size. Increasing the sample size will increase the power of the statistical test. This is usually the recommended method to ensure sufficient statistical power.

Controlling Extraneous Variables. Controlling for extraneous variables reduces the population variance, so the statistical test is more powerful. Controlling for extraneous variable(s) does change the hypothesis being tested.

Comparing the Extremes of the Population or Comparing Extremes of the Independent Variable can increase the statistical power…given there are differences in the population. The drawback is that the results will not generalize to the entire population.

One Sample *t* Test

A researcher wanted to know whether students receiving academic scholarships are significantly different in age than the general MTSU student population—which has an average age of 24. A sample of 25 students receiving academic scholarships reported an average age of 21 with a standard deviation of 3 years. Use an alpha of .05 to test the null hypothesis.

1. **State the null and alternative hypotheses.** Usually the hypotheses concern the value of a population parameter such as the population mean.

Null (H_0):

Alternative (H_1):

2. **Set the criteria for a decision.** Before we select a sample, we use the null hypothesis to predict the characteristics that the sample should have.

   Alpha (α) =

   Degrees of freedom (*df*) = *n – 1* =

   Critical Values:

3. **Collect data and compute sample statistics.**

Hypothesized population mean (μ) =

Sample size (*n*) =

Sample mean (*M*) =

Sum of squares (*SS*) = *df*variance*

Sample variance (s^2) = *SS/df* =

Sample SD (*s*) =

estimated standard error between M and $\mu = \frac{SD}{\sqrt{n}} = \sqrt{\frac{\text{Variance}}{\text{Sample Size}}} = \sqrt{\frac{s^2}{n}} = s_M$

$$t = \frac{\text{sample mean – hypothesized population mean}}{\text{estimated standard error between } M \text{ and } \mu} = \frac{M - \mu}{s_M}$$

$$\text{estimated Cohen's } d = \frac{\text{sample mean-hypothesized population mean}}{\text{sample standard deviation}} = \frac{M - \mu}{s} = \hat{d}$$

$$eta^2 = \frac{t^2}{t^2 + df}$$

4. **Make a decision.** We compare the obtained sample data with the prediction that was made from the null hypothesis. If the sample mean is consistent with the prediction, we conclude the null hypothesis is reasonable. But if there is a big discrepancy between the data and the prediction, we decide that the null hypothesis is wrong.

   t **<u>was NOT in critical region (retain null because *p* > alpha).</u>** Using an alpha of .05, the one sample *t* test indicated the average GPA for the sorority members (M = 3.02, SD = 0.40, n = 43) was not significantly different than the GPA of females within the general student population (μ = 2.98), t (42) = 1.89, $\hat{d}$ = 0.10, $p > .05$.

   t **<u>was in critical region (reject null because *p* < alpha).</u>** Using an alpha of .05, the one sample *t* test indicated the average GPA for the sorority members (M = 3.02, SD = 0.75, n = 43) was significantly different than the GPA of females within the general student population (μ = 2.80), t (42) = 2.77, $\hat{d}$ = .29, $p < .05$.

A researcher wanted to estimate the difference in average loneliness between college students and the general population (μ = 40). Fifty college students were given a loneliness survey; their average score was found to be 42 with a standard deviation of nine points. Calculate and interpret the 95% confidence interval.

Determine the critical values.

- ❑ Two-tailed critical values must be used (i.e., analogous to nondirectional hypotheses).
- ❑ Use alpha (α) = .05 for a 95% confidence interval.
- ❑ Use alpha (α) = .01 for a 99% confidence interval.

$df = n - 1$ **Critical Values:**

Identify the information in the scenario.

Population mean (μ) = Sum of squares (SS) = df*$variance$

Sample size (n) = Sample variance (s^2) = SS/df =

Sample mean (M) = Sample SD (s) =

Calculate the Confidence Interval

estimated standard error between M and $\mu = \frac{SD}{\sqrt{n}} = \sqrt{\frac{\text{Variance}}{\text{Sample Size}}} = \sqrt{\frac{s^2}{n}} = s_M$

Lower Limit
$(M - \mu) - t_{crit}(s_M)$

Upper Limit
$(M - \mu) + t_{crit}(s_M)$

Interpret the Confidence Interval
We are 95% or 99% confident that the interval (lower limit, upper limit) covers the mean difference in dependent variable between the sample (M = __, SD = __ , n =__) and the population (μ = __).

A faculty member wanted to know how the stress levels of graduate students compares to the general population. The faculty member administered a stress test to the graduate students. The stress test is known to have an average of 40 for the general population. Use an alpha of .05 to test the null hypothesis.

SPSS Instructions
→Analyze→Compare Means→One-Sample T Test
Test Variable(s): Stress
Test Value: 40

T-Test

One-Sample Statistics

	N	Mean	Std. Deviation	Std. Error Mean
stress	17	50.1765	6.69174	1.62299

One-Sample Test

	Test Value = 40					
					95% Confidence Interval of the Difference	
	t	df	Sig. (2-tailed)	Mean Difference	Lower	Upper
stress	6.270	16	.000	10.17647	6.7359	13.6170

1. **State the null and alternative hypotheses.** Usually the hypotheses concern the value of a population parameter such as the population mean.

Null (H_0):

Alternative (H_1):

2. **Set the criteria for a decision.** Before we select a sample, we use the null hypothesis to predict the characteristics that the sample should have.

   Alpha (α) =

3. **Collect data and compute sample statistics.**

$$\text{estimated Cohen's } d = \frac{\text{sample mean-hypothesized population mean}}{\text{sample standard deviation}} = \frac{M - \mu}{s} = \hat{d}$$

$$eta^2 = \frac{t^2}{t^2 + df}$$

4. **Make a decision.** We compare the obtained sample data with the prediction that was made from the null hypothesis. If the sample mean is consistent with the prediction, we conclude the null hypothesis is reasonable. But if there is a big discrepancy between the data and the prediction, we decide that the null hypothesis is wrong.

   Identify the *p* value.
   - *p equals the sig. value if it was a nondirectional hypothesis.*
   - *p equals the sig. value/2 if it was a directional hypothesis. The results also would have to match the predicted direction.*

   p =

   Retain null because *p* > alpha. Using an alpha of .05, the one sample *t* test indicated the average GPA for the sorority members ($M = 3.02$, $SD = 0.40$, $n = 43$) was not significantly different than the GPA of females within the general student population ($\mu = 2.98$), $t(42) = 1.89$, $\hat{d} = 0.10$, $p = .073$.

   Reject null because *p* < alpha. Using an alpha of .05, the one sample *t* test indicated the average GPA for the sorority members ($M = 3.02$, $SD = 0.75$, $n = 43$) was significantly different than the GPA of females within the general student population ($\mu = 2.80$), $t(42) = 2.77$, $\hat{d} = .29$, $p = .002$.

Interpret the Confidence Interval

Assumptions for the One-Sample *t* Test

Independent Observations
The individuals being measured are assumed to contribute one and only one score and not be influenced by anyone else in the study. The method used to collect the data determines whether this is a valid assumption.

> It is very important that the independence assumption be valid. Otherwise, the test results may be dramatically misleading….empirical alpha does not match nominal alpha; power is different than expected.

Normal Sampling Distribution
We assume the population is normal; if our sample size is 30 or larger, the sampling distribution will be approximately normal.

> Violating the normality assumption does not have a major impact on the test results as long as the sample size is above 30.

t Test for Two Independent Samples

A psychologist wanted to determine whether the local residence for students (on-campus, off-campus) is related to alcohol consumption as measured by number of drinks per week. The psychologist surveyed students enrolled in a statistics class and found that the 40 on-campus students consumed an average of 5.5 drinks per week with a standard deviation of 2.5 drinks. The 40 off-campus students consumed an average of 4.8 drinks with a standard deviation of 1.5 drinks. Use an alpha of .01 to test the null hypothesis.

1. **State the null and alternative hypotheses.** Usually the hypotheses concern the value population parameters such as the population mean.

Null (H_0):

Alternative (H_1):

2. **Set the criteria for a decision.** Before we select a sample, we use the null hypothesis to predict the characteristics that the sample should have.

   Alpha (α) =

   Degrees of freedom (df) =

   Critical Values:

3. **Collect data and compute sample statistics.**

Sample 1	Sample 2
Sample size (n_1) = Mean (M_1) = Sum of squares (SS_1) = df_1*$variance_1$ Variance ($s^2{}_1$) = SS_1/df_1 = SD (s_1) =	Sample size (n_2) = Mean (M_2) = Sum of squares (SS_2) = df_2*$variance_2$ Variance ($s^2{}_2$) = SS_2/df_2 = SD (s_2) =

$$Pooled\ Variance = \frac{SS_1 + SS_2}{df_1 + df_2} = s_P^2$$

$$Estimated\ Standard\ Error = \sqrt{\frac{Pooled\ Variance}{n_1} + \frac{Pooled\ Variance}{n_2}} = \sqrt{\frac{s_P^2}{n_1} + \frac{s_P^2}{n_2}}$$
$$= s_{(M_1 - M_2)}$$

$$t = \frac{(SampleMean_1 - SampleMean_2) - (0)}{estimated\ standard\ error\ of (M_1 - M_2)} = \frac{M_1 - M_2}{s_{(M_1 - M_2)}}$$

$$\text{estimated Cohen's } d = \frac{SampleMean_1\text{-}SampleMean_2}{\sqrt{Pooled\ Variance}} = \frac{M_1 - M_2}{s_P} = \hat{d}$$

$$eta^2 = \frac{t^2}{t^2 + df}$$

4. **Make a decision.** We compare the obtained sample data with the prediction that was made from the null hypothesis. If the sample data are consistent with the prediction, we conclude the null hypothesis is reasonable. But if there is a big discrepancy between the data and the prediction, we decide that the null hypothesis is wrong.

   <u>*t* was NOT in critical region (retain null because *p* > alpha).</u> Using an alpha of .05, the independent samples *t* test indicated the average GPA for the sorority members (*M* = 3.02, *SD* = 0.40, *n* = 43) was not significantly different than the average GPA for fraternity members (*M = 2.98, SD* = 0.50, *n* = 43), $t\ (84) = 1.89$, $\hat{d} = 0.10$, $p > .05$.

   <u>*t* was in critical region (reject null because *p* < alpha).</u> Using an alpha of .05, the independent samples *t* test indicated the average GPA for the sorority members (*M* = 2.93, *SD* = 0.40, *n* = 43) was significantly different than the average GPA for fraternity members (*M = 2.58, SD* = 0.50, *n* = 43), $t\ (84) = 4.12$, $\hat{d} = 0.10$, $p < .05$.

Confidence Intervals for the Independent Samples *t* Situation

A researcher wanted to estimate the difference in average anxiety between females and males. Twenty-eight females reported an average anxiety of 41 (*SD* = 5); twenty-eight males reported an average anxiety of 38 (*SD* = 7). Calculate and interpret the 95% confidence interval.

Determine the critical values.

- ❑ Two-tailed critical values must be used (i.e., analogous to nondirectional hypotheses).
- ❑ Use alpha (α) = .05 for a 95% confidence interval.
- ❑ Use alpha (α) = .01 for a 99% confidence interval.

$df = n_1 + n_2 - 2$ **Critical Values:**

Identify the information in the scenario.

Sample 1	Sample 2
Sample size (n_1) = Mean (M_1) = Sum of squares (SS_1) = df_1*$variance_1$ Variance (s^2_1) = SS_1/df_1 = SD (s_1) =	Sample size (n_2) = Mean (M_2) = Sum of squares (SS_2) = df_2*$variance_2$ Variance (s^2_2) = SS_2/df_2 = SD (s_2) =

Calculate the Confidence Interval

$$Pooled\ Variance = \frac{SS_1 + SS_2}{df_1 + df_2} = s_P^2$$

$$Estimated\ Standard\ Error = \sqrt{\frac{Pooled\ Variance}{n_1} + \frac{Pooled\ Variance}{n_2}} = \sqrt{\frac{s_P^2}{n_1} + \frac{s_P^2}{n_2}}$$

$$= s_{(M_1 - M_2)}$$

Lower Limit

$(M_1 - M_2) - t_{crit}(s_{(M_1 - M_2)})$

Upper Limit

$(M_1 - M_2) + t_{crit}(s_{(M_1 - M_2)})$

Interpret the Confidence Interval

We are <u>95% or 99%</u> confident that the interval (lower limit, upper limit) covers the mean difference in <u>dependent variable</u> between <u>sample1</u> (*M* = __, *SD* = __ , *n* =__) and <u>sample 2</u> (*M* = __, *SD* = __ , *n* =__).

A developmental psychologist wanted to determine whether girls and boys have the same number of imaginary friends during childhood. The psychologist asked children at a local elementary school about the number of imaginary friends they have (and confirmed this with their caregiver). Use an alpha of .05 to test the null hypothesis.

SPSS Instructions
→Analyze→Compare Means→Independent Samples T Test
Test Variable(s): ImaginaryFriends
Grouping Variable: Gender
→Define Groups
Group 1: 1
Group 2: 2

T-Test

Group Statistics

	GENDER	N	Mean	Std. Deviation	Std. Error Mean
Imaginary Friends	Girls	15	2.0000	1.77281	.45774
	Boys	15	3.0000	1.88982	.48795

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
Imaginary Friends	Equal variances assumed	.106	.747	-1.495	28	.146	-1.0000	.66904	-2.37047	.37047
	Equal variances not assumed			-1.495	27.886	.146	-1.0000	.66904	-2.37072	.37072

1. **State the null and alternative hypotheses.** Usually the hypotheses concern the value of a population parameter such as the population mean.

Null (H_0):

Alternative (H_1):

2. **Set the criteria for a decision.** Before we select a sample, we use the null hypothesis to predict the characteristics that the sample should have.

   Alpha (α) =

3. Compute effect size estimates.

$$\text{estimated Cohen's } d = \frac{SampleMean_1\text{-}SampleMean_2}{\sqrt{Pooled\ Variance}} = \frac{M_1 - M_2}{s_P} = \hat{d}$$

$$eta^2 = \frac{t^2}{t^2 + df}$$

4. Make a decision. We compare the obtained sample data with the prediction that was made from the null hypothesis. If the sample mean is consistent with the prediction, we conclude the null hypothesis is reasonable. But if there is a big discrepancy between the data and the prediction, we decide that the null hypothesis is wrong.

Identify the *p* value.

- ❑ *p equals the sig. value if it was a nondirectional hypothesis.*
- ❑ *p equals the sig. value/2 if it was a directional hypothesis. The results also would have to match the predicted direction.*

p =

Retain null because *p* > alpha. Using an alpha of .05, the independent samples *t* test indicated the average GPA for the sorority members (*M* = 3.02, *SD* = 0.40, *n* = 43) was not significantly different than the average GPA for fraternity members (*M = 2.98, SD* = 0.50, *n* = 43), *t* (84) = 1.89, $\hat{d}$ = 0.10, *p* = .24.

Reject null because *p* < alpha. Using an alpha of .05, the independent samples *t* test indicated the average GPA for the sorority members (*M* = 2.93, *SD* = 0.40, *n* = 43) was significantly different than the average GPA for fraternity members (*M = 2.58, SD* = 0.50, *n* = 43), *t* (84) = 4.12, $\hat{d}$ = 0.10, *p* = .021.

Interpret the Confidence Interval

We are 95% or 99% confident that the interval (lower limit, upper limit) covers the mean difference in dependent variable between sample1 (*M* = __, *SD* = __ , *n* =__) and sample 2 (*M* = __, *SD* = __ , *n* =__).

Assumptions for the *t* Test for Independent Samples

Independent Observations within Each Sample
The individuals being measured are assumed to contribute one and only one score and not be influenced by anyone else in the study. The method used to collect the data determines whether this is a valid assumption.

> It is very important that the independence assumption be valid. Otherwise, the test results may be dramatically misleading.

Normal Sampling Distribution for Each Population
We assume the population is normal; if our sample size is 30 or larger, the sampling distribution will be approximately normal.

> Violating the normality assumption does not have a major impact on the test results as long as the sample size is above 30.

Equal Variances for the Two Populations (i.e., homogeneity of variance)
The formula is based on a pooled variance. It makes sense to average these two variances only if they both estimate the same population variance.

Assumptions for the *Welch t* Test for Independent Samples

Independent Observations within Each Sample
The individuals being measured are assumed to contribute one and only one score and not be influenced by anyone else in the study. The method used to collect the data determines whether this is a valid assumption.

> It is very important that the independence assumption be valid. Otherwise, the test results may be dramatically misleading.

Normal Sampling Distribution for Each Population
We assume the population is normal; if our sample size is 30 or larger, the sampling distribution will be approximately normal.

> Violating the normality assumption does not have a major impact on the test results as long as the sample size is above 30.

Choosing between the *t* Test for Independent Samples and the Welch *t* Test

Welch t Test for Two Independent Samples

A researcher compared the salaries of faculty at universities to faculty at junior colleges. Thirty university faculty members reported an average salary of 60 thousand dollars with a variance of 21 thousand dollars squared. Twenty-four faculty members working at junior colleges reported an average salary of 54 thousand dollars with a variance of 12 thousand dollars squared. Use an alpha of .05 to test the null hypothesis.

1. **State the null and alternative hypotheses.** Usually the hypotheses concern the value population parameters such as the population mean.

Null (H_0):

Alternative (H_1):

2. **Set the criteria for a decision.** Before we select a sample, we use the null hypothesis to predict the characteristics that the sample should have.

   Alpha (α) =

   $$Welch\ df = \frac{(V_1 + V_2)^2}{\frac{V_1^2}{n_1 - 1} + \frac{V_2^2}{n_2 - 1}} \text{ where } V_1 = \frac{s_1^2}{n_1} \text{ and } V_2 = \frac{s_2^2}{n_2}$$

   Critical Values:

3. **Collect data and compute sample statistics.**

Sample 1	Sample 2
Sample size (n_1) = Mean (M_1) = Sum of squares (SS_1) = df_1*$variance_1$ Variance ($s^2{}_1$) = SS_1/df_1 = SD (s_1) =	Sample size (n_2) = Mean (M_2) = Sum of squares (SS_2) = df_2*$variance_2$ Variance ($s^2{}_2$) = SS_2/df_2 = SD (s_2) =

$$Estimated\ Standard\ Error = \sqrt{\frac{Variance_1}{n_1} + \frac{Variance_2}{n_2}} = \sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}} = s_{(M_1 - M_2)}$$

$$Welch\ t = \frac{(SampleMean_1 - SampleMean_2) - (0)}{estimated\ standard\ error\ of (M_1 - M_2)} = \frac{M_1 - M_2}{s_{(M_1 - M_2)}}$$

$$\text{estimated Cohen's } d = \frac{SampleMean_1\text{-}SampleMean_2}{\sqrt{\frac{\text{Variance}_1 + \text{Variance}_2}{2}}} = \frac{M_1 - M_2}{\sqrt{\frac{s_1^2 + s_2^2}{2}}} = \hat{d}$$

$$eta^2 = \frac{t^2}{t^2 + df}$$

4. **Make a decision.** We compare the obtained sample data with the prediction that was made from the null hypothesis. If the sample data are consistent with the prediction, we conclude the null hypothesis is reasonable. But if there is a big discrepancy between the data and the prediction, we decide that the null hypothesis is wrong.

   ***t* was NOT in critical region (retain null because *p* > alpha).** Using an alpha of .05, the Welch *t* test for independent samples indicated the average GPA for the sorority members (*M* = 3.02, *SD* = 0.40, *n* = 40) was not significantly different than the average GPA for fraternity members (*M = 2.98, SD* = 0.50, *n* = 23), *t* (58.7) = 1.89, $\hat{d}$ = 0.10, *p* > .05.

   ***t* was in critical region (reject null because *p* < alpha).** Using an alpha of .05, the Welch *t* test for independent samples indicated the average GPA for the sorority members (*M* = 2.93, *SD* = 0.40, *n* = 40) was significantly different than the average GPA for fraternity members (*M = 2.58, SD* = 0.50, *n* = 23), *t* (58.7) = 4.12, $\hat{d}$ = 0.24, *p* < .05.

A researcher wanted to estimate the difference in average salary between university and junior college faculty. Thirty university faculty members reported an average salary of 60 thousand dollars with a variance of 21 thousand dollars squared. Twenty-four faculty members working at junior colleges reported an average salary of 54 thousand dollars with a variance of 12 thousand dollars squared. Calculate and interpret the 95% confidence interval.

Determine the critical values.

- Two-tailed critical values must be used (i.e., analogous to nondirectional hypotheses).
- Use alpha (α) = .05 for a 95% confidence interval.
- Use alpha (α) = .01 for a 99% confidence interval.

$$Welch\ df = \frac{(V_1 + V_2)^2}{\frac{V_1^2}{n_1 - 1} + \frac{V_2^2}{n_2 - 1}} \text{ where } V_1 = \frac{s_1^2}{n_1} \text{ and } V_2 = \frac{s_2^2}{n_2}$$

$$V_1 = \frac{21}{30} = .7 \text{ and } V_2 = \frac{12}{24} = .5 \text{ and } Welch\ df = \frac{(.7+.5)^2}{\frac{.7^2}{30-1}+\frac{.5^2}{24-1}} = \frac{1.44}{.0169+.0109} = \mathbf{51.8}$$

Critical Values:

Sample 1	Sample 2
Sample size (n_1) = Mean (M_1) = Sum of squares (SS_1) = df_1*$variance_1$ Variance (s^2_1) = SS_1/df_1 = SD (s_1) =	Sample size (n_2) = Mean (M_2) = Sum of squares (SS_2) = df_2*$variance_2$ Variance (s^2_2) = SS_2/df_2 = SD (s_2) =

Calculate the Confidence Interval

$$Estimated\ Standard\ Error = \sqrt{\frac{Variance_1}{n_1} + \frac{Variance_2}{n_2}} = \sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}} = s_{(M_1 - M_2)}$$

Lower Limit
$(M_1 - M_2) - t_{crit}(s_{(M_1-M_2)})$

Upper Limit
$(M_1 - M_2) + t_{crit}(s_{(M_1-M_2)})$

Interpret the Confidence Interval
We are 95% or 99% confident that the interval (lower limit, upper limit) covers the mean difference in dependent variable between sample1 (*M* = __, *SD* = __ , *n* =__) and sample 2 (*M* = __, *SD* = __ , *n* =__).

A researcher wanted to know whether individuals in an opposite-sex marriage have similar satisfaction levels as individuals in a same-sex marriage. The researcher contacted 22 individuals asked them to rate their satisfaction level. Use an alpha of .05 to test the null hypothesis.

SPSS Instructions
→Analyze→Compare Means→Independent Samples T Test
Test Variable(s): Satisfaction
Grouping Variable: Marriage
→Define Groups
Group 1: 1
Group 2: 2

T-Test

Group Statistics

	marriage	N	Mean	Std. Deviation	Std. Error Mean
satisfaction	Same-Sex	11	9.4545	1.91644	.57783
	Opposite-Sex	11	8.5455	4.00908	1.20878

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
satisfaction	Equal variances assumed	3.422	.079	.679	20	.505	.90909	1.33979	-1.88567	3.70385
	Equal variances not assumed			.679	14.343	.508	.90909	1.33979	-1.95804	3.77622

1. **State the null and alternative hypotheses.** Usually the hypotheses concern the value of a population parameter such as the population mean.

Null (H_0):

Alternative (H_1):

2. **Set the criteria for a decision.** Before we select a sample, we use the null hypothesis to predict the characteristics that the sample should have.

   Alpha (α) =

3. Compute effect size estimates.

$$\text{estimated Cohen's } d = \frac{SampleMean_1\text{-}SampleMean_2}{\sqrt{\frac{\text{Variance}_1 + \text{Variance}_2}{2}}} = \frac{M_1 - M_2}{\sqrt{\frac{s_1^2 + s_2^2}{2}}} = \hat{d}$$

$$eta^2 = \frac{t^2}{t^2 + df}$$

4. Make a decision. We compare the obtained sample data with the prediction that was made from the null hypothesis. If the sample mean is consistent with the prediction, we conclude the null hypothesis is reasonable. But if there is a big discrepancy between the data and the prediction, we decide that the null hypothesis is wrong.

Identify the *p* value.

- *p equals the sig. value if it was a nondirectional hypothesis.*
- *p equals the sig. value/2 if it was a directional hypothesis. The results also would have to match the predicted direction.*

p =

<u>Retain null because *p* > alpha.</u> Using an alpha of .05, the Welch *t* test for independent samples indicated the average GPA for the sorority members (*M* = 3.02, *SD* = 0.40, *n* = 10) was not significantly different than the average GPA for fraternity members (*M = 2.98, SD* = 0.50, *n* = 10), *t* (17.4) = 1.89, $\hat{d}$ = 0.10, *p* = .24.

<u>Reject null because *p* < alpha.</u> Using an alpha of .05, the Welch *t* test for independent samples indicated the average GPA for the sorority members (*M* = 2.93, *SD* = 0.40, *n* = 40) was significantly different than the average GPA for fraternity members (*M = 2.58, SD* = 0.50, *n* = 23), *t* (47.2) = 4.12, $\hat{d}$ = 0.10, *p* = .021.

Interpret the Confidence Interval

We are <u>95% or 99%</u> confident that the interval (lower limit, upper limit) covers the mean difference in <u>dependent variable</u> between <u>sample1 </u>(*M* = __, *SD* = __ , *n* =__) and <u>sample 2</u> (*M* = __, *SD* = __ , *n* =__).

t Test for Two Dependent Samples (a.k.a., *t* Test for Related Samples)

There are three research designs that would lead to dependence between the two samples.

1. You have two scores for each person.

2. You have naturally occurring pairs of people (e.g., couples, parent-child).

3. You control for an extraneous variable by matching subjects on the extraneous variable.

A researcher wanted to compare the average aggression levels of parents and their children. The researcher randomly selected 10 households that had children in them and administered aggression scales to the primary care giver and the oldest child in the home. Use an alpha of .05 to test the null hypothesis.

	Parent	Child		**Difference = Parent - Child**
Aggression	15	12		
Level	17	15		
	17	16		
	11	10		
	14	9		
	16	18		
	18	10		
	26	19		
	15	10		
	15	15		
Mean	16.4	13.4	Sample mean (M_D) =	
SD	3.8930	3.6576	Sample SD (s_D) =	3.1972
			Sample size (n_D) =	10

1. **State the null and alternative hypotheses.** Usually the hypotheses concern the value of a population parameter such as the population mean.

   Null (H_0):

   Alternative (H_1):

2. **Set the criteria for a decision.** Before we select a sample, we use the null hypothesis to predict the characteristics that the sample should have.

   Alpha (α) =

   Degrees of freedom (*df*) = n_D – 1

   Critical Values:

3. **Collect data and compute sample statistics.**

   Hypothesized population mean (μ_D) =

   Sample size (n_D) =

   Sample mean (M_D) =

   Sum of squares (SS_D) = *df*variance*

   Sample variance (s_D^2) = SS_D/df =

   Sample SD (s_D) =

$$\text{estimated standard error between } M_D \text{ and } \mu_D = \sqrt{\frac{Variance_{\text{D}}}{Sample\ Size_{\text{D}}}} = \sqrt{\frac{s_D^2}{n_D}} = s_{M_D}$$

$$t = \frac{\text{sample mean}_{\text{D}} - \text{hypothesized population mean}_{\text{D}}}{\text{estimated standard error between M}_{\text{D}} \text{ and } \mu_{\text{D}}} = \frac{M_D - \mu_D}{s_{M_D}}$$

4. Compute effect size estimates.

$$\text{estimated Cohen's } d = \frac{\text{sample mean-hypothesized population mean}}{\text{sample standard deviation}} = \frac{M_D - \mu_D}{s_D} = \hat{d}$$

$$eta^2 = \frac{t^2}{t^2 + df}$$

5. **Make a decision.** We compare the obtained sample data with the prediction that was made from the null hypothesis. If the sample mean is consistent with the prediction, we conclude the null hypothesis is reasonable. But if there is a big discrepancy between the data and the prediction, we decide that the null hypothesis is wrong.

   <u>*t* was NOT in critical region (retain null because *p* > alpha).</u> Using an alpha of .05, the related samples *t* test indicated the average hostility for the parents (M = 3.68, SD = 1.08, n = 16) was not significantly different than the average hostility for the children (M = 3.18, SD = 2.19, n = 16), $t\ (30)$ = 1.89, $\hat{d}$ = .11, p > .05.

   <u>*t* was in critical region (reject null because *p* < alpha).</u> Using an alpha of .05, the related samples *t* test indicated the average hostility for the parents (M = 3.68, SD = 1.08, n = 16) was significantly different than the average hostility for the children (M = 3.18, SD = 2.19, n = 16), $t\ (30)$ = 2.89, $\hat{d}$ = .31, p < .05.

Confidence Intervals for the Related Samples *t* Situation

A researcher wanted to estimate the difference in chemistry knowledge from the first day to the last day of the semester. The researcher gave a comprehensive exam on the first day of the semester (*M* = 35, *SD* = 12, *n* = 20) and again on the last day of the semester (*M* = 75, *SD* = 15, *n* = 20). The average difference was 40 points (*SD* = 18). Calculate and interpret the 95% confidence interval for the difference.

Determine the critical values.

- Use the mean, SD, and sample size of the difference scores.
- Two-tailed critical values must be used (i.e., analogous to nondirectional hypotheses).
- Use alpha (α) = .05 for a 95% confidence interval.
- Use alpha (α) = .01 for a 99% confidence interval.

$df = n_D - 1$ **Critical Values:**

Identify the information in the scenario.

Hypothesized population mean (μ_D) = Sum of squares (SS_D) = *df*variance*

Sample size (n_D) = Sample variance (s_D^2) = SS_D/df =

Sample mean (M_D) = Sample SD (s_D) =

Calculate the Confidence Interval

$$\text{estimated standard error between } M_D \text{ and } \mu_D = \sqrt{\frac{Variance_{\mathrm{D}}}{Sample\ Size_{\mathrm{D}}}} = \sqrt{\frac{s_D^2}{n_D}} = s_{M_D}$$

Lower Limit

$(M_D - \mu_D) - t_{crit}(s_M)$

Upper Limit

$(M_D - \mu_D) - t_{crit}(s_M)$

Interpret the Confidence Interval

We are <u>95% or 99%</u> confident that the interval (lower limit, upper limit) covers the mean difference in <u>dependent variable</u> between <u>sample1</u> (*M* = __, *SD* = __ , *n* =__) and <u>sample 2</u> (*M* = __, *SD* = __ , *n* =__).

A statistics professor asked her students to rate (1 = extremely negative, 10 = extremely positive) their attitudes about statistics at the start of the semester and at the end of the semester. To evaluate whether the teacher had an impact on students attitudes about statistics, she compared their ratings.Use an alpha of .05 to test the null hypothesis.

SPSS Instructions
→Analyze→Compare Means→Paired-Samples T Test
Variable 1: AttitudeStart
Variable 2: AttitudeEnd

T-Test

Paired Samples Statistics

		Mean	N	Std. Deviation	Std. Error Mean
Pair 1	attitude @ start	4.7460	63	2.02380	.25497
	attitude @ end	5.9683	63	1.75947	.22167

Paired Samples Correlations

		N	Correlation	Sig.
Pair 1	attitude @ start & attitude @ end	63	.451	.000

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	attitude @ start - attitude @ end	-1.22222	1.99551	.25141	-1.72479	-.71966	-4.861	62	.000

1. **State the null and alternative hypotheses.** Usually the hypotheses concern the value of a population parameter such as the population mean.

Null (H_0):

Alternative (H_1):

2. **Set the criteria for a decision.** Before we select a sample, we use the null hypothesis to predict the characteristics that the sample should have.

   Alpha (α) =

3. Compute effect size estimates.

$$\text{estimated Cohen's } d = \frac{\text{sample mean-hypothesized population mean}}{\text{sample standard deviation}} = \frac{M_D - \mu_D}{s_D} = \hat{d}$$

$$eta^2 = \frac{t^2}{t^2 + df}$$

4. Make a decision. We compare the obtained sample data with the prediction that was made from the null hypothesis. If the sample mean is consistent with the prediction, we conclude the null hypothesis is reasonable. But if there is a big discrepancy between the data and the prediction, we decide that the null hypothesis is wrong.

Identify the *p* value.

- *p equals the sig. value if it was a nondirectional hypothesis.*
- *p equals the sig. value/2 if it was a directional hypothesis. The results also would have to match the predicted direction.*

p **=**

Retain null because *p* > alpha. Using an alpha of .05, the related samples *t* test indicated the average hostility for the parents (M = 3.68, SD = 1.08, n = 16) was not significantly different than the average hostility for the children (M = 3.18, SD = 2.19, n = 16), $t\,(30)$ = 1.89, $\hat{d}$ = .11, p = .34.

Reject null because *p* < alpha. Using an alpha of .05, the related samples *t* test indicated the average hostility for the parents (M = 3.68, SD = 1.08, n = 16) was significantly different than the average hostility for the children (M = 3.18, SD = 2.19, n = 16), $t\,(30)$ = 2.89, $\hat{d}$ = .31, p = .012.

Interpret the Confidence Interval

We are 95% or 99% confident that the interval (lower limit, upper limit) covers the mean difference in dependent variable between sample1 (*M* = __, *SD* = __ , *n* =__) and sample 2 (*M* = __, *SD* = __ , *n* =__).

Assumptions of the Related Samples *t* Test

Independent Observations

The difference scores are assumed to be independent of one another. In a within-subjects design, it means each person is not influenced by any other person. In a matched-subjects design, it means each pair is not influenced by any other pair. The method used to collect the data determines whether this is a valid assumption.

> It is very important that the independence assumption be valid. Otherwise, the test results may be dramatically misleading.

Normal Sampling Distribution
We assume the population is normal for the difference scores; if our sample size is 30 or larger, the sampling distribution will be approximately normal.

> Violating the normality assumption does not have a major impact on the test results as long as the sample size is above 30.

CRITICAL VALUES of the "t" Distribution

	Proportion in ONE Tail					
	.25	.10	.05	.025	.01	.005
	Proportion in TWO Tails					
DF	.50	.20	.10	.05	.02	.01
1	1.00000	3.07768	6.31375	12.70620	31.82052	63.65674
2	0.81650	1.88562	2.91999	4.30265	6.96456	9.92484
3	0.76489	1.63774	2.35336	3.18245	4.54070	5.84091
4	0.74070	1.53321	2.13185	2.77645	3.74695	4.60409
5	0.72669	1.47588	2.01505	2.57058	3.36493	4.03214
6	0.71756	1.43976	1.94318	2.44691	3.14267	3.70743
7	0.71114	1.41492	1.89458	2.36462	2.99795	3.49948
8	0.70639	1.39682	1.85955	2.30600	2.89646	3.35539
9	0.70272	1.38303	1.83311	2.26216	2.82144	3.24984
10	0.69981	1.37218	1.81246	2.22814	2.76377	3.16927
11	0.69745	1.36343	1.79588	2.20099	2.71808	3.10581
12	0.69548	1.35622	1.78229	2.17881	2.68100	3.05454
13	0.69383	1.35017	1.77093	2.16037	2.65031	3.01228
14	0.69242	1.34503	1.76131	2.14479	2.62449	2.97684
15	0.69120	1.34061	1.75305	2.13145	2.60248	2.94671
16	0.69013	1.33676	1.74588	2.11991	2.58349	2.92078
17	0.68920	1.33338	1.73961	2.10982	2.56693	2.89823
18	0.68836	1.33039	1.73406	2.10092	2.55238	2.87844
19	0.68762	1.32773	1.72913	2.09302	2.53948	2.86093
20	0.68695	1.32534	1.72472	2.08596	2.52798	2.84534
21	0.68635	1.32319	1.72074	2.07961	2.51765	2.83136
22	0.68581	1.32124	1.71714	2.07387	2.50832	2.81876
23	0.68531	1.31946	1.71387	2.06866	2.49987	2.80734
24	0.68485	1.31784	1.71088	2.06390	2.49216	2.79694
25	0.68443	1.31635	1.70814	2.05954	2.48511	2.78744
26	0.68404	1.31497	1.70562	2.05553	2.47863	2.77871
27	0.68368	1.31370	1.70329	2.05183	2.47266	2.77068
28	0.68335	1.31253	1.70113	2.04841	2.46714	2.76326
29	0.68304	1.31143	1.69913	2.04523	2.46202	2.75639
30	0.68276	1.31042	1.69726	2.04227	2.45726	2.75000
31	0.68249	1.30946	1.69552	2.03951	2.45282	2.74404
32	0.68223	1.30857	1.69389	2.03693	2.44868	2.73848
33	0.68200	1.30774	1.69236	2.03452	2.44479	2.73328
34	0.68177	1.30695	1.69092	2.03224	2.44115	2.72839
35	0.68156	1.30621	1.68957	2.03011	2.43772	2.72381
36	0.68137	1.30551	1.68830	2.02809	2.43449	2.71948
37	0.68118	1.30485	1.68709	2.02619	2.43145	2.71541
38	0.68100	1.30423	1.68595	2.02439	2.42857	2.71156
39	0.68083	1.30364	1.68488	2.02269	2.42584	2.70791
40	0.68067	1.30308	1.68385	2.02108	2.42326	2.70446
41	0.68052	1.30254	1.68288	2.01954	2.42080	2.70118
42	0.68038	1.30204	1.68195	2.01808	2.41847	2.69807
43	0.68024	1.30155	1.68107	2.01669	2.41625	2.69510
44	0.68011	1.30109	1.68023	2.01537	2.41413	2.69228
45	0.67998	1.30065	1.67943	2.01410	2.41212	2.68959
46	0.67986	1.30023	1.67866	2.01290	2.41019	2.68701
47	0.67975	1.29982	1.67793	2.01174	2.40835	2.68456
48	0.67964	1.29944	1.67722	2.01063	2.40658	2.68220
49	0.67953	1.29907	1.67655	2.00958	2.40489	2.67995
50	0.67943	1.29871	1.67591	2.00856	2.40327	2.67779

CRITICAL VALUES of the "t" Distribution

	Proportion in ONE Tail					
	.25	.10	.05	.025	.01	.005
	Proportion in TWO Tails					
DF	.50	.20	.10	.05	.02	.01
51	0.67933	1.29837	1.67528	2.00758	2.40172	2.67572
52	0.67924	1.29805	1.67469	2.00665	2.40022	2.67373
53	0.67915	1.29773	1.67412	2.00575	2.39879	2.67182
54	0.67906	1.29743	1.67356	2.00488	2.39741	2.66998
55	0.67898	1.29713	1.67303	2.00404	2.39608	2.66822
56	0.67890	1.29685	1.67252	2.00324	2.39480	2.66651
57	0.67882	1.29658	1.67203	2.00247	2.39357	2.66487
58	0.67874	1.29632	1.67155	2.00172	2.39238	2.66329
59	0.67867	1.29607	1.67109	2.00100	2.39123	2.66176
60	0.67860	1.29582	1.67065	2.00030	2.39012	2.66028
61	0.67853	1.29558	1.67022	1.99962	2.38905	2.65886
62	0.67847	1.29536	1.66980	1.99897	2.38801	2.65748
63	0.67840	1.29513	1.66940	1.99834	2.38701	2.65615
64	0.67834	1.29492	1.66901	1.99773	2.38604	2.65485
65	0.67828	1.29471	1.66864	1.99714	2.38510	2.65360
66	0.67823	1.29451	1.66827	1.99656	2.38419	2.65239
67	0.67817	1.29432	1.66792	1.99601	2.38330	2.65122
68	0.67811	1.29413	1.66757	1.99547	2.38245	2.65008
69	0.67806	1.29394	1.66724	1.99495	2.38161	2.64898
70	0.67801	1.29376	1.66691	1.99444	2.38081	2.64790
71	0.67796	1.29359	1.66660	1.99394	2.38002	2.64686
72	0.67791	1.29342	1.66629	1.99346	2.37926	2.64585
73	0.67787	1.29326	1.66600	1.99300	2.37852	2.64487
74	0.67782	1.29310	1.66571	1.99254	2.37780	2.64391
75	0.67778	1.29294	1.66543	1.99210	2.37710	2.64298
76	0.67773	1.29279	1.66515	1.99167	2.37642	2.64208
77	0.67769	1.29264	1.66488	1.99125	2.37576	2.64120
78	0.67765	1.29250	1.66462	1.99085	2.37511	2.64034
79	0.67761	1.29236	1.66437	1.99045	2.37448	2.63950
80	0.67757	1.29222	1.66412	1.99006	2.37387	2.63869
81	0.67753	1.29209	1.66388	1.98969	2.37327	2.63790
82	0.67749	1.29196	1.66365	1.98932	2.37269	2.63712
83	0.67746	1.29183	1.66342	1.98896	2.37212	2.63637
84	0.67742	1.29171	1.66320	1.98861	2.37156	2.63563
85	0.67739	1.29159	1.66298	1.98827	2.37102	2.63491
86	0.67735	1.29147	1.66277	1.98793	2.37049	2.63421
87	0.67732	1.29136	1.66256	1.98761	2.36998	2.63353
88	0.67729	1.29125	1.66235	1.98729	2.36947	2.63286
89	0.67726	1.29114	1.66216	1.98698	2.36898	2.63220
90	0.67723	1.29103	1.66196	1.98667	2.36850	2.63157
91	0.67720	1.29092	1.66177	1.98638	2.36803	2.63094
92	0.67717	1.29082	1.66159	1.98609	2.36757	2.63033
93	0.67714	1.29072	1.66140	1.98580	2.36712	2.62973
94	0.67711	1.29062	1.66123	1.98552	2.36667	2.62915
95	0.67708	1.29053	1.66105	1.98525	2.36624	2.62858
96	0.67705	1.29043	1.66088	1.98498	2.36582	2.62802
97	0.67703	1.29034	1.66071	1.98472	2.36541	2.62747
98	0.67700	1.29025	1.66055	1.98447	2.36500	2.62693
99	0.67698	1.29016	1.66039	1.98422	2.36461	2.62641
100	0.67695	1.29007	1.66023	1.98397	2.36422	2.62589