Module 8: Introduction to the *t* tests

Jon Starkweather, PhD
jonathan.starkweather@unt.edu
Consultant
Research and Statistical Support

UNT UNIVERSITY OF NORTH TEXAS
Discover the power of ideas.

Introduction to Statistics for the Social Sciences

RSS
Research and Statistical Support

The RSS short courses

The Research and Statistical Support (RSS) office at the University of North Texas hosts a number of “Short Courses”. A list of them is available at:

`http://www.unt.edu/rss/Instructional.htm`

Outline

Outline

- One Sample *t* Test

Outline

- One Sample *t* Test
  - The *t* Distribution
  - One Sample *t* Test NHST Example
  - Effect Size (Cohen's *d*)
  - Using delta (δ) for Statistical Power
  - Calculating Confidence Interval Limits (CI_{95})
  - Summary of Section 1

RSS
Research and Statistical Support

Outline

- One Sample *t* Test
  - The *t* Distribution
  - One Sample *t* Test NHST Example
  - Effect Size (Cohen's *d*)
  - Using delta (δ) for Statistical Power
  - Calculating Confidence Interval Limits (CI_{95})
  - Summary of Section 1
- Dependent Samples *t* Test

RSS
Research and Statistical Support

Outline

- One Sample t Test
  - The t Distribution
  - One Sample t Test NHST Example
  - Effect Size (Cohen's d)
  - Using delta (δ) for Statistical Power
  - Calculating Confidence Interval Limits (CI_{95})
  - Summary of Section 1
- Dependent Samples t Test
  - The Dependent Samples t Test
  - Hypothesis Testing Example
  - Effect Size (Cohen's d)
  - Using delta (δ) for Statistical Power
  - Calculating Confidence Interval Limits (CI_{95})
  - Summary of Section 2

RSS
Research and Statistical Support

Outline

- Continued on the next slide.

RSS
Research and Statistical Support

Outline continued

Outline continued

- Independent Samples t Test

Outline continued

- Independent Samples *t* Test
  - The Independent Samples *t* Test
  - Hypothesis Testing Example
  - Effect Size (Cohen's *d*)
  - Using delta (δ) for Statistical Power
  - Calculating Confidence Interval Limits (CI_{95})
  - Summary of Section 3

RSS
Research and Statistical Support

Outline continued

- Independent Samples t Test
  - The Independent Samples t Test
  - Hypothesis Testing Example
  - Effect Size (Cohen's d)
  - Using delta (δ) for Statistical Power
  - Calculating Confidence Interval Limits (CI_{95})
  - Summary of Section 3
- Summary of t Tests

RSS
Research and Statistical Support

Outline continued

Outline continued

RSS
Research and Statistical Support

One Sample *t* test

- An extension of the Z-test with a sample mean.

One Sample *t* test

- An extension of the Z-test with a sample mean.
  - Still comparing a sample mean to a population mean.

RSS
Research and Statistical Support

One Sample *t* test

- An extension of the Z-test with a sample mean.
  - Still comparing a sample mean to a population mean.
- But, we have the problem of an unknown σ and σ^2.

One Sample *t* test

- An extension of the Z-test with a sample mean.
  - Still comparing a sample mean to a population mean.
- But, we have the problem of an unknown σ and σ^2.
- Both S and S^2 are biased because samples are by their very nature, smaller than the population(s) they represent.

Research and Statistical Support

One Sample *t* test

- An extension of the Z-test with a sample mean.
  - Still comparing a sample mean to a population mean.
- But, we have the problem of an unknown σ and σ^2.
- Both S and S^2 are biased because samples are by their very nature, smaller than the population(s) they represent.
- So, we apply a *progressive correction* $(n - 1)$, to arrive at an **unbiased estimate of the population variance**.

One Sample *t* test

- An extension of the Z-test with a sample mean.
  - Still comparing a sample mean to a population mean.
- But, we have the problem of an unknown σ and σ^2.
- Both S and S^2 are biased because samples are by their very nature, smaller than the population(s) they represent.
- So, we apply a *progressive correction* ($n - 1$), to arrive at an **unbiased estimate of the population variance**.

$$S^2 = \frac{\sum(X - \overline{X})^2}{n - 1}$$

One Sample *t* test

- An extension of the Z-test with a sample mean.
  - Still comparing a sample mean to a population mean.
- But, we have the problem of an unknown σ and σ^2.
- Both S and S^2 are biased because samples are by their very nature, smaller than the population(s) they represent.
- So, we apply a *progressive correction* $(n-1)$, to arrive at an **unbiased estimate of the population variance**.

$$S^2 = \frac{\sum(X-\overline{X})^2}{n-1}$$

- We say *progressive correction* because; as sample size (n) gets larger, the correction has *less* effect.

RSS
Research and Statistical Support

One Sample *t* test

- An extension of the Z-test with a sample mean.
  - Still comparing a sample mean to a population mean.
- But, we have the problem of an unknown σ and σ^2.
- Both S and S^2 are biased because samples are by their very nature, smaller than the population(s) they represent.
- So, we apply a *progressive correction* $(n-1)$, to arrive at an **unbiased estimate of the population variance.**

$$S^2 = \frac{\sum(X-\overline{X})^2}{n-1}$$

- We say *progressive correction* because; as sample size (n) gets larger, the correction has *less* effect.
  - Larger sample $\rightarrow$ includes more of the population $\rightarrow$ less bias.

RSS
Research and Statistical Support

One Sample t test

- An extension of the Z-test with a sample mean.
  - Still comparing a sample mean to a population mean.
- But, we have the problem of an unknown σ and σ^2.
- Both S and S^2 are biased because samples are by their very nature, smaller than the population(s) they represent.
- So, we apply a *progressive correction* ($n - 1$), to arrive at an **unbiased estimate of the population variance**.

$$S^2 = \frac{\sum(X - \overline{X})^2}{n-1}$$

- We say *progressive correction* because; as sample size (n) gets larger, the correction has *less* effect.
  - Larger sample $\rightarrow$ includes more of the population $\rightarrow$ less bias.
  - Therefore, less correction is needed as sample size increases.

RSS
Research and Statistical Support

Degrees of Freedom (*df*)

We call this progressive correction the **degrees of Freedom** (*df*).

Degrees of Freedom (*df*)

We call this progressive correction the **degrees of Freedom** (*df*).

- The number of scores in a sample which are "free to vary".

RSS
Research and Statistical Support

Degrees of Freedom (*df*)

We call this progressive correction the **degrees of Freedom** (*df*).

- The number of scores in a sample which are "free to vary".
  - If you know the mean and all but one of the scores, you can figure the score you don't know.

RSS
Research and Statistical Support

Degrees of Freedom (*df*)

We call this progressive correction the **degrees of Freedom (*df*).**

- The number of scores in a sample which are “free to vary”.
  - If you know the mean and all but one of the scores, you can figure the score you don’t know.
  - So, for one sample *t* test: $df = n - 1$

RSS
Research and Statistical Support

Degrees of Freedom (*df*)

We call this progressive correction the **degrees of Freedom** (*df*).

- The number of scores in a sample which are "free to vary".
  - If you know the mean and all but one of the scores, you can figure the score you don't know.
  - So, for one sample *t* test: $df = n - 1$
- Now we can go on with some of the 'usual stuff'.

RSS
Research and Statistical Support

Degrees of Freedom (*df*)

We call this progressive correction the **degrees of Freedom** (*df*).

- The number of scores in a sample which are "free to vary".
  - If you know the mean and all but one of the scores, you can figure the score you don't know.
  - So, for one sample *t* test: $df = n - 1$
- Now we can go on with some of the 'usual stuff'.
  - Distribution of Means...

Research and Statistical Support

Degrees of Freedom (*df*)

We call this progressive correction the **degrees of Freedom** (*df*).

- The number of scores in a sample which are "free to vary".
  - If you know the mean and all but one of the scores, you can figure the score you don't know.
  - So, for one sample *t* test: $df = n - 1$
- Now we can go on with some of the 'usual stuff'.
  - Distribution of Means...

$$S_M^2 = \frac{S^2}{n} \qquad S_M = \sqrt{S_M^2}$$

RSS
Research and Statistical Support

Comparison Distribution Changes

- Because we are estimating the population variance, we use the *t* distribution (instead of the *Z* distribution as was done previously).

RSS
Research and Statistical Support

Comparison Distribution Changes

- Because we are estimating the population variance, we use the *t* distribution (instead of the *Z* distribution as was done previously).
  - The *t* distribution is **not** normally distributed, due to more error in estimation (i.e., estimating the unknown σ^2).

RSS
Research and Statistical Support

Comparison Distribution Changes

- Because we are estimating the population variance, we use the *t* distribution (instead of the *Z* distribution as was done previously).
  - The *t* distribution is **not** normally distributed, due to more error in estimation (i.e., estimating the unknown σ^2).
- Different table, similar procedure for finding the cutoff sample score associated with $p < .05$, or any other significance level.

RSS Research and Statistical Support

Comparison Distribution Changes

- Because we are estimating the population variance, we use the *t* distribution (instead of the *Z* distribution as was done previously).
  - The *t* distribution is **not** normally distributed, due to more error in estimation (i.e., estimating the unknown σ^2).
- Different table, similar procedure for finding the cutoff sample score associated with $p < .05$, or any other significance level.
  - But, you must take into account sample size; meaning, the *df* in order to find t_{crit}

RSS Research and Statistical Support

An Excerpt from the *t* distribution table

Table 1: *t* Critical Values

1-tailed	*df*	.10	.05	.025	.01
2-tailed	*df*	.20	.10	.05	.02
	1	3.078	6.314	12.710	31.821
	2	1.886	2.920	4.303	6.965
	3	1.638	2.353	3.182	4.541
	4	1.533	2.132	2.776	3.747
	5	1.476	2.015	2.571	3.365
	6	1.440	1.943	2.447	3.143
	etc.				

Degrees of freedom in the left column and significance level along the top rows.

RSS
Research and Statistical Support

Calculating the One Sample t

- Also referred to as t_{calc} which is compared to t_{crit}

RSS
Research and Statistical Support

Calculating the One Sample t

- Also referred to as t_{calc} which is compared to t_{crit}
- Recall, S_M is the standard deviation of a distribution of means.

Calculating the One Sample t

- Also referred to as t_{calc} which is compared to t_{crit}
- Recall, S_M is the standard deviation of a distribution of means.
- First, calculate S_M (μ is given):

RSS
Research and Statistical Support

Calculating the One Sample t

- Also referred to as t_{calc} which is compared to t_{crit}
- Recall, S_M is the standard deviation of a distribution of means.
- First, calculate S_M (μ is given): $S^2 = \frac{\sum(X - \overline{X})^2}{n-1}$ then

RSS Research and Statistical Support

Calculating the One Sample t

- Also referred to as t_{calc} which is compared to t_{crit}
- Recall, S_M is the standard deviation of a distribution of means.
- First, calculate S_M (μ is given): $S^2 = \frac{\sum(X - \overline{X})^2}{n-1}$ then $S_M^2 = \frac{S^2}{n}$ then

Research and Statistical Support

Calculating the One Sample t

- Also referred to as t_{calc} which is compared to t_{crit}
- Recall, S_M is the standard deviation of a distribution of means.
- First, calculate S_M (μ is given): $S^2 = \frac{\sum(X - \overline{X})^2}{n-1}$ then
  $S_M^2 = \frac{S^2}{n}$ then
  $S_M = \sqrt{S_M^2}$

RSS Research and Statistical Support

Calculating the One Sample t

- Also referred to as t_{calc} which is compared to t_{crit}
- Recall, S_M is the standard deviation of a distribution of means.
- First, calculate S_M (μ is given): $S^2 = \frac{\sum (X - \overline{X})^2}{n-1}$ then
  $S_M^2 = \frac{S^2}{n}$ then
  $S_M = \sqrt{S_M^2}$
- Which leads to:

RSS Research and Statistical Support

Calculating the One Sample t

- Also referred to as t_{calc} which is compared to t_{crit}
- Recall, S_M is the standard deviation of a distribution of means.
- First, calculate S_M (μ is given): $S^2 = \frac{\sum(X-\overline{X})^2}{n-1}$ then
  $S_M^2 = \frac{S^2}{n}$ then
  $S_M = \sqrt{S_M^2}$
- Which leads to:

$$t_{calc} = \frac{\overline{X}-\mu}{S_M} = \frac{\overline{X}-\mu}{S/\sqrt{n}} = \frac{\overline{X}-\mu}{\sqrt{S^2/n}}$$

RSS Research and Statistical Support

Calculating the One Sample t

- Also referred to as t_{calc} which is compared to t_{crit}
- Recall, S_M is the standard deviation of a distribution of means.
- First, calculate S_M (μ is given): $S^2 = \frac{\sum(X - \overline{X})^2}{n-1}$ then
  $S_M^2 = \frac{S^2}{n}$ then
  $S_M = \sqrt{S_M^2}$
- Which leads to:

$$t_{calc} = \frac{\overline{X} - \mu}{S_M} = \frac{\overline{X} - \mu}{S/\sqrt{n}} = \frac{\overline{X} - \mu}{\sqrt{S^2/n}}$$

  - The formula below is the one used here but, all three above are equivalent (t_{calc}).

Calculating the One Sample t

- Also referred to as t_{calc} which is compared to t_{crit}
- Recall, S_M is the standard deviation of a distribution of means.
- First, calculate S_M (μ is given): $S^2 = \frac{\sum(X - \overline{X})^2}{n-1}$ then
  $S_M^2 = \frac{S^2}{n}$ then
  $S_M = \sqrt{S_M^2}$
- Which leads to:

$$t_{calc} = \frac{\overline{X} - \mu}{S_M} = \frac{\overline{X} - \mu}{S/\sqrt{n}} = \frac{\overline{X} - \mu}{\sqrt{S^2/n}}$$

  - The formula below is the one used here but, all three above are equivalent (t_{calc}).

$$t_{calc} = \frac{\overline{X} - \mu}{S_M}$$

RSS
Research and Statistical Support

NHST Example

Research question:

NHST Example

Research question:

- Does this class of four students get **less** sleep than all UNT students ($\mu = 6.08$ hours)?

RSS
Research and Statistical Support

NHST Example

Research question:

- Does this class of four students get **less** sleep than all UNT students ($\mu = 6.08$ hours)?
- **Step 1:** Define the populations and restate the research question as null and alternative hypotheses.

RSS
Research and Statistical Support

NHST Example

Research question:

- Does this class of four students get **less** sleep than all UNT students ($\mu = 6.08$ hours)?
- **Step 1:** Define the populations and restate the research question as null and alternative hypotheses.
- Population 1: Students in this class.

RSS
Research and Statistical Support

NHST Example

Research question:

- Does this class of four students get **less** sleep than all UNT students ($\mu = 6.08$ hours)?
- **Step 1:** Define the populations and restate the research question as null and alternative hypotheses.
- Population 1: Students in this class.
- Population 2: All other UNT students.

NHST Example

Research question:

- Does this class of four students get **less** sleep than all UNT students ($\mu = 6.08$ hours)?
- **Step 1:** Define the populations and restate the research question as null and alternative hypotheses.
- Population 1: Students in this class.
- Population 2: All other UNT students.

$$H_0 : \mu_1 \geq \mu_2$$
$$H_1 : \mu_1 < \mu_2$$

RSS Research and Statistical Support

NHST Example

Research question:

- Does this class of four students get **less** sleep than all UNT students ($\mu = 6.08$ hours)?
- **Step 1:** Define the populations and restate the research question as null and alternative hypotheses.
- Population 1: Students in this class.
- Population 2: All other UNT students.

$$H_0 : \mu_1 \geq \mu_2$$
$$H_1 : \mu_1 < \mu_2$$

- Note, the Alternative hypothesis (H_1) is directional: one-tailed test.

RSS
Research and Statistical Support

Step 2: Comparison Distribution

Table 2: Comparison dist. (estimate σ with S_M)

X	$\overline{X}$	$X - \overline{X}$	$\left(X - \overline{X}\right)^2$
5	4	1	1
3	4	-1	1
6	4	2	4
2	4	-2	4
16			$SOS = 10$

RSS
Research and Statistical Support

Step 2: Comparison Distribution

Table 2: Comparison dist. (estimate σ with S_M)

X	$\overline{X}$	$X - \overline{X}$	$\left(X - \overline{X}\right)^2$
5	4	1	1
3	4	-1	1
6	4	2	4
2	4	-2	4
16			$SOS = 10$

$$S^2 = \frac{\sum (x - \overline{x})^2}{n-1} = \frac{10}{4-1} = 3.33$$

RSS
Research and Statistical Support

Step 2: Comparison Distribution

Table 2: Comparison dist. (estimate σ with S_M)

X	$\overline{X}$	$X-\overline{X}$	$\left(X-\overline{X}\right)^2$
5	4	1	1
3	4	-1	1
6	4	2	4
2	4	-2	4
16			$SOS = 10$

$$S^2 = \frac{\sum\left(x-\overline{x}\right)^2}{n-1} = \frac{10}{4-1} = 3.33$$

$$S_M^2 = \frac{S^2}{n} = \frac{3.33}{4} = 0.8325$$

RSS
Research and Statistical Support

Step 2: Comparison Distribution

Table 2: Comparison dist. (estimate σ with S_M)

X	$\overline{X}$	$X - \overline{X}$	$\left(X - \overline{X}\right)^2$
5	4	1	1
3	4	-1	1
6	4	2	4
2	4	-2	4
16			$SOS = 10$

$$S^2 = \frac{\sum\left(x - \overline{x}\right)^2}{n-1} = \frac{10}{4-1} = 3.33$$

$$S_M^2 = \frac{S^2}{n} = \frac{3.33}{4} = 0.8325$$

$$S_M = \sqrt{S_M^2} = \sqrt{0.8325} = 0.912$$

Step 3: Determine the critical score

- Determine the cutoff sample score.

Step 3: Determine the critical score

- Determine the cutoff sample score.
- Significance level = .05, one-tailed test, $df = 4 - 1 = 3$

Step 3: Determine the critical score

- Determine the cutoff sample score.
- Significance level = .05, one-tailed test, $df = 4 - 1 = 3$
- Look in the appropriate column of the *t* table.

RSS
Research and Statistical Support

Step 3: Determine the critical score

- Determine the cutoff sample score.
- Significance level = .05, one-tailed test, $df = 4 - 1 = 3$
- Look in the appropriate column of the *t* table.
  - One-tailed, significance level = .05

`http://www.math.unb.ca/~knight/utility/t-table.htm`

RSS Research and Statistical Support

Step 3: Determine the critical score

- Determine the cutoff sample score.
- Significance level = .05, one-tailed test, $df = 4 - 1 = 3$
- Look in the appropriate column of the *t* table.
    - One-tailed, significance level = .05

`http://www.math.unb.ca/~knight/utility/t-table.htm`

- Critical *t* value is 2.353

RSS
Research and Statistical Support

Step 3: Determine the critical score

- Determine the cutoff sample score.
- Significance level = .05, one-tailed test, $df = 4 - 1 = 3$
- Look in the appropriate column of the *t* table.
  - One-tailed, significance level = .05

`http://www.math.unb.ca/~knight/utility/t-table.htm`

- Critical *t* value is 2.353
- Which means, we need t_{calc} to be greater than $|2.353|$ (absolute value of 2.353) to find a significant difference.

RSS
Research and Statistical Support

Step 4: Calculate t (t_{calc})

- From previous calculating, we have:

Step 4: Calculate t (t_{calc})

- From previous calculating, we have:

$$\overline{X} = 4.00,\ S_M = 0.912$$

Step 4: Calculate t (t_{calc})

- From previous calculating, we have:

$$\overline{X} = 4.00,\ S_M = 0.912$$

- So, we can plug in the values, including the population mean which was given.

RSS
Research and Statistical Support

Step 4: Calculate t (t_{calc})

- From previous calculating, we have:

$$\overline{X} = 4.00, \; S_M = 0.912$$

- So, we can plug in the values, including the population mean which was given.

$$t_{calc} = \frac{\overline{X} - \mu}{S_M} = \frac{4.00 - 6.08}{0.912} = \frac{-2.08}{0.912} = -2.28$$

RSS
Research and Statistical Support

Step 5: Compare and Make a Decision

- Since $t_{calc} = |-2.28| < |2.353| = t_{crit}$ we fail to reject the null hypothesis.

RSS
Research and Statistical Support

Step 5: Compare and Make a Decision

- Since $t_{calc} = |-2.28| < |2.353| = t_{crit}$ we fail to reject the null hypothesis.
- The interpretation is...Students in this class ($M = 4.00$) do not sleep significantly fewer hours than all other students at UNT ($\mu = 6.08$), $t(3) = -2.28, p > .05$.

Step 5: Compare and Make a Decision

- Since $t_{calc} = |-2.28| < |2.353| = t_{crit}$ we fail to reject the null hypothesis.
- The interpretation is...Students in this class ($M = 4.00$) do not sleep significantly fewer hours than all other students at UNT ($\mu = 6.08$), $t(3) = -2.28, p > .05$.
- Do not be tempted to say something like; nearly significant, just missed significance, etc.

RSS
Research and Statistical Support

Step 5: Compare and Make a Decision

- Since $t_{calc} = |-2.28| < |2.353| = t_{crit}$ we fail to reject the null hypothesis.
- The interpretation is...Students in this class ($M = 4.00$) do not sleep significantly fewer hours than all other students at UNT ($\mu = 6.08$), $t(3) = -2.28, p > .05$.
- Do not be tempted to say something like; nearly significant, just missed significance, etc.
    - Although the sample mean is numerically smaller, it is not statistically significantly smaller.

RSS
Research and Statistical Support

Step 5: Compare and Make a Decision

- Since $t_{calc} = |-2.28| < |2.353| = t_{crit}$ we fail to reject the null hypothesis.
- The interpretation is...Students in this class ($M = 4.00$) do not sleep significantly fewer hours than all other students at UNT ($\mu = 6.08$), $t(3) = -2.28$, $p > .05$.
- Do not be tempted to say something like; nearly significant, just missed significance, etc.
  - Although the sample mean is numerically smaller, it is not statistically significantly smaller.
  - Remember, we have a *very* small sample size, so we would need a *very* large mean difference to achieve significance (or a larger sample).

RSS
Research and Statistical Support

Effect Size

- The appropriate effect size measure for the one sample *t* test is Cohen's *d*.

Effect Size

- The appropriate effect size measure for the one sample t test is Cohen's d.
- Calculation of d in its general form (as it was with the Z-test) is:

Research and Statistical Support

Effect Size

- The appropriate effect size measure for the one sample *t* test is Cohen's *d*.
- Calculation of *d* in its general form (as it was with the Z-test) is:

$$d = \frac{\overline{X} - \mu}{\sigma}$$

Effect Size

- The appropriate effect size measure for the one sample *t* test is Cohen's *d*.
- Calculation of *d* in its general form (as it was with the Z-test) is:

$$d = \frac{\overline{X} - \mu}{\sigma}$$

- However, we do not know the population standard deviation (σ) in the *t* situation, so we estimate with $S = \sqrt{S^2}$

Effect Size

- The appropriate effect size measure for the one sample *t* test is Cohen's *d*.
- Calculation of *d* in its general form (as it was with the Z-test) is:

$$d = \frac{\overline{X} - \mu}{\sigma}$$

- However, we do not know the population standard deviation (σ) in the *t* situation, so we estimate with $S = \sqrt{S^2}$

$$d = \frac{\overline{X} - \mu}{S} = \frac{4 - 6.08}{\sqrt{3.33}} = \frac{-2.08}{1.82} = 1.143$$

RSS Research and Statistical Support

Effect Size

- The appropriate effect size measure for the one sample *t* test is Cohen's *d*.
- Calculation of *d* in its general form (as it was with the Z-test) is:

$$d = \frac{\overline{X} - \mu}{\sigma}$$

- However, we do not know the population standard deviation (σ) in the *t* situation, so we estimate with $S = \sqrt{S^2}$

$$d = \frac{\overline{X} - \mu}{S} = \frac{4 - 6.08}{\sqrt{3.33}} = \frac{-2.08}{1.82} = 1.143$$

- So, although we have a large effect size (standardized difference), we did not achieve statistical significance. However, keep in mind that with a larger sample, this amount of mean difference may have been significant.

RSS
Research and Statistical Support

Effect Size

- The appropriate effect size measure for the one sample *t* test is Cohen's *d*.
- Calculation of *d* in its general form (as it was with the Z-test) is:

$$d = \frac{\overline{X} - \mu}{\sigma}$$

- However, we do not know the population standard deviation (σ) in the *t* situation, so we estimate with $S = \sqrt{S^2}$

$$d = \frac{\overline{X} - \mu}{S} = \frac{4 - 6.08}{\sqrt{3.33}} = \frac{-2.08}{1.82} = 1.143$$

- So, although we have a large effect size (standardized difference), we did not achieve statistical significance. However, keep in mind that with a larger sample, this amount of mean difference may have been significant.
  - Statistical significance is **directly** tied to sample size, effect size is not.

RSS
Research and Statistical Support

Statistical Power and Sample Size

- Calculating power with the one sample t test is slightly different than with the Z-test.

Statistical Power and Sample Size

- Calculating power with the one sample t test is slightly different than with the Z-test.
- Here, we introduce something called the *Non-Centrality Parameter* (NCP).

Statistical Power and Sample Size

- Calculating power with the one sample *t* test is slightly different than with the Z-test.
- Here, we introduce something called the *Non-Centrality Parameter* (NCP).
  - The NCP takes its name from the fact that if H_0 is false, the distribution of our statistic will not be centered on zero, but on something called delta.

RSS
Research and Statistical Support

Statistical Power and Sample Size

- Calculating power with the one sample *t* test is slightly different than with the Z-test.
- Here, we introduce something called the *Non-Centrality Parameter* (NCP).
  - The NCP takes its name from the fact that if H_0 is false, the distribution of our statistic will not be centered on zero, but on something called delta.
    - Thus, the distribution is *non-central*.

Statistical Power and Sample Size

- Calculating power with the one sample *t* test is slightly different than with the Z-test.
- Here, we introduce something called the *Non-Centrality Parameter* (NCP).
  - The NCP takes its name from the fact that if H_0 is false, the distribution of our statistic will not be centered on zero, but on something called delta.
    - Thus, the distribution is *non-central*.
    - The symbol for delta is: δ

RSS
Research and Statistical Support

Statistical Power and Sample Size

- Calculating power with the one sample *t* test is slightly different than with the Z-test.
- Here, we introduce something called the *Non-Centrality Parameter* (NCP).
  - The NCP takes its name from the fact that if H_0 is false, the distribution of our statistic will not be centered on zero, but on something called delta.
    - Thus, the distribution is *non-central*.
    - The symbol for delta is: δ
- In the context of the one sample *t* test, delta is calculated as follows:

RSS Research and Statistical Support

Statistical Power and Sample Size

- Calculating power with the one sample *t* test is slightly different than with the Z-test.
- Here, we introduce something called the *Non-Centrality Parameter* (NCP).
  - The NCP takes its name from the fact that if H_0 is false, the distribution of our statistic will not be centered on zero, but on something called delta.
    - Thus, the distribution is *non-central*.
    - The symbol for delta is: δ
- In the context of the one sample *t* test, delta is calculated as follows:

$$\delta = d * \sqrt{n}$$

RSS
Research and Statistical Support

Statistical Power and Sample Size

- To calculate delta (δ) for our current example:

Statistical Power and Sample Size

- To calculate delta (δ) for our current example:

$$\delta = 1.143 * \sqrt{4} = 1.143 * 2 = 2.286$$

RSS
Research and Statistical Support

Statistical Power and Sample Size

- To calculate delta (δ) for our current example:

$$\delta = 1.143 * \sqrt{4} = 1.143 * 2 = 2.286$$

- With the significance level (.05, one-tailed), we can look up power in a δ table.

Statistical Power and Sample Size

- To calculate delta (δ) for our current example:

$$\delta = 1.143 * \sqrt{4} = 1.143 * 2 = 2.286$$

- With the significance level (.05, one-tailed), we can look up power in a δ table.

  http://www.unt.edu/rss/class/Jon/ISSS_SC/Module008/delta

Statistical Power and Sample Size

- To calculate delta (δ) for our current example:

$$\delta = 1.143 * \sqrt{4} = 1.143 * 2 = 2.286$$

- With the significance level (.05, one-tailed), we can look up power in a δ table.

  http://www.unt.edu/rss/class/Jon/ISSS_SC/Module008/delta

- Power for this example was $\approx$ 0.89 which is not terribly meaningful once the study has been completed.

RSS
Research and Statistical Support

Statistical Power and Sample Size

- To calculate delta (δ) for our current example:

$$\delta = 1.143 * \sqrt{4} = 1.143 * 2 = 2.286$$

- With the significance level (.05, one-tailed), we can look up power in a δ table.

  `http://www.unt.edu/rss/class/Jon/ISSS_SC/Module008/delta`

- Power for this example was $\approx$ 0.89 which is not terribly meaningful once the study has been completed.
- However, we can now use δ to calculate sample size a-priori for a given power (prior to collecting data for the next study).

RSS
Research and Statistical Support

Statistical Power and Sample Size

- To calculate delta (δ) for our current example:

$$\delta = 1.143 * \sqrt{4} = 1.143 * 2 = 2.286$$

- With the significance level (.05, one-tailed), we can look up power in a δ table.

  http://www.unt.edu/rss/class/Jon/ISSS_SC/Module008/delta

- Power for this example was $\approx$ 0.89 which is not terribly meaningful once the study has been completed.
- However, we can now use δ to calculate sample size a-priori for a given power (prior to collecting data for the next study).
  - Where $n = (\delta/d)^2$
  - This type of power calculation (figuring a-priori sample size) is **very** useful.

RSS
Research and Statistical Support

Calculating a Confidence Interval

- Confidence intervals allow us to go beyond NHST and offer us an idea of the location of a population mean (μ).

Calculating a Confidence Interval

- Confidence intervals allow us to go beyond NHST and offer us an idea of the location of a population mean (μ).
- We need three numbers to calculate a confidence interval:

RSS
Research and Statistical Support

Calculating a Confidence Interval

- Confidence intervals allow us to go beyond NHST and offer us an idea of the location of a population mean (μ).
- We need three numbers to calculate a confidence interval:
  - The Standard Error (SE) which is also known as the standard deviation of a distribution of means: $S_M = 0.912$

RSS Research and Statistical Support

Calculating a Confidence Interval

- Confidence intervals allow us to go beyond NHST and offer us an idea of the location of a population mean (μ).
- We need three numbers to calculate a confidence interval:
  - The Standard Error (SE) which is also known as the standard deviation of a distribution of means: $S_M = 0.912$
  - The critical value from our current study: $t_{crit} = 2.353$

RSS
Research and Statistical Support

Calculating a Confidence Interval

- Confidence intervals allow us to go beyond NHST and offer us an idea of the location of a population mean (μ).
- We need three numbers to calculate a confidence interval:
  - The Standard Error (SE) which is also known as the standard deviation of a distribution of means: $S_M = 0.912$
  - The critical value from our current study: $t_{crit} = 2.353$
  - And the mean of our sample: $\overline{X} = 4.00$

RSS
Research and Statistical Support

Calculating a Confidence Interval

- Confidence intervals allow us to go beyond NHST and offer us an idea of the location of a population mean (μ).
- We need three numbers to calculate a confidence interval:
  - The Standard Error (SE) which is also known as the standard deviation of a distribution of means: $S_M = 0.912$
  - The critical value from our current study: $t_{crit} = 2.353$
  - And the mean of our sample: $\overline{X} = 4.00$
- Recall the general formulas for a confidence interval from Module 6:

Calculating a Confidence Interval

- Confidence intervals allow us to go beyond NHST and offer us an idea of the location of a population mean (μ).
- We need three numbers to calculate a confidence interval:
  - The Standard Error (SE) which is also known as the standard deviation of a distribution of means: $S_M = 0.912$
  - The critical value from our current study: $t_{crit} = 2.353$
  - And the mean of our sample: $\overline{X} = 4.00$
- Recall the general formulas for a confidence interval from Module 6:

$$UpperLimit(UL) = (+criticalvalue) * (SE) + \overline{X}$$
$$LowerLimit(LL) = (-criticalvalue) * (SE) + \overline{X}$$

RSS
Research and Statistical Support

Calculating a CI

- In this example, we are calculating a 95% confidence interval (CI_{95}) because our critical value was based on a significance level of .05 (1 - .05 = .95 or 95%).

RSS
Research and Statistical Support

Calculating a CI

- In this example, we are calculating a 95% confidence interval (CI_{95}) because our critical value was based on a significance level of .05 (1 - .05 = .95 or 95%).
- $t_{crit} = 2.353$, $S_M = 0.912$, $\overline{X} = 4.00$

Calculating a CI

- In this example, we are calculating a 95% confidence interval (CI_{95}) because our critical value was based on a significance level of .05 (1 - .05 = .95 or 95%).
- $t_{crit} = 2.353$, $S_M = 0.912$, $\overline{X} = 4.00$

$$UL = +2.353 * 0.912 + 4.00 = 6.146$$
$$LL = -2.353 * 0.912 + 4.00 = 1.854$$

RSS
Research and Statistical Support

Calculating a CI

- In this example, we are calculating a 95% confidence interval (CI_{95}) because our critical value was based on a significance level of .05 (1 - .05 = .95 or 95%).
- $t_{crit} = 2.353$, $S_M = 0.912$, $\overline{X} = 4.00$

$$UL = +2.353 * 0.912 + 4.00 = 6.146$$
$$LL = -2.353 * 0.912 + 4.00 = 1.854$$

- If we were to take an infinite number of samples of students in this class, 95% of those samples' means would be between 6.146 and 1.854 hours of sleep.

RSS
Research and Statistical Support

Calculating a CI

- In this example, we are calculating a 95% confidence interval (CI_{95}) because our critical value was based on a significance level of .05 (1 - .05 = .95 or 95%).
- $t_{crit} = 2.353$, $S_M = 0.912$, $\overline{X} = 4.00$

$$UL = +2.353 * 0.912 + 4.00 = 6.146$$
$$LL = -2.353 * 0.912 + 4.00 = 1.854$$

- If we were to take an infinite number of samples of students in this class, 95% of those samples' means would be between 6.146 and 1.854 hours of sleep.
  - Remember, the population mean is fixed (but unknown); while each sample has its own mean (sample means fluctuate).

RSS
Research and Statistical Support

CI_{95} Considerations and Interpretations

The current example resulted in $CI_{95} = 6.146 : 1.854$.

CI_{95} Considerations and Interpretations

The current example resulted in $CI_{95} = 6.146 : 1.854$.

- Notice the population 2 mean ($\mu_2 = 6.08$), representing all UNT students, falls **inside** our interval.

CI_{95} Considerations and Interpretations

The current example resulted in $CI_{95} = 6.146 : 1.854$.

- Notice the population 2 mean ($\mu_2 = 6.08$), representing all UNT students, falls **inside** our interval.
  - Recall, we did not reject the null; meaning our sample *did* come from population 2, not a distinct population (i.e. population 1).

RSS
Research and Statistical Support

CI_{95} Considerations and Interpretations

The current example resulted in $CI_{95} = 6.146 : 1.854$.

- Notice the population 2 mean ($\mu_2 = 6.08$), representing all UNT students, falls **inside** our interval.
  - Recall, we did not reject the null; meaning our sample *did* come from population 2, not a distinct population (i.e. population 1).
- If μ_2 were greater than 6.146, then we would have rejected the null hypothesis and inferred that our sample came from population 1; meaning, population 1 would have been significantly different from population 2.

RSS
Research and Statistical Support

CI_{95} Considerations and Interpretations

The current example resulted in $CI_{95} = 6.146 : 1.854$.

- Notice the population 2 mean ($\mu_2 = 6.08$), representing all UNT students, falls **inside** our interval.
  - Recall, we did not reject the null; meaning our sample *did* come from population 2, not a distinct population (i.e. population 1).
- If μ_2 were greater than 6.146, then we would have rejected the null hypothesis and inferred that our sample came from population 1; meaning, population 1 would have been significantly different from population 2.
- Important: the interval is **not** interpreted as "we are 95% confident that population 1's mean is between 6.146 and 1.854."

CI_{95} Considerations and Interpretations

The current example resulted in $CI_{95} = 6.146 : 1.854$.

- Notice the population 2 mean ($\mu_2 = 6.08$), representing all UNT students, falls **inside** our interval.
  - Recall, we did not reject the null; meaning our sample *did* come from population 2, not a distinct population (i.e. population 1).
- If μ_2 were greater than 6.146, then we would have rejected the null hypothesis and inferred that our sample came from population 1; meaning, population 1 would have been significantly different from population 2.
- Important: the interval is **not** interpreted as "we are 95% confident that population 1's mean is between 6.146 and 1.854."
  - We are dealing with a sample; we do not know what μ_1 is; so, we can not know what that interval would be.

One Sample *t* test Usage?

Unfortunately...

One Sample *t* test Usage?

Unfortunately...

- The one sample t test is essentially never used, but it servers a good purpose to familiarize us with the *t* distribution.

RSS
Research and Statistical Support

One Sample *t* test Usage?

Unfortunately...

- The one sample t test is essentially never used, but it servers a good purpose to familiarize us with the *t* distribution.
- Occasionally, we know both μ and σ, for instance with SAT or GRE scores, which necessitates the Z-test.

RSS
Research and Statistical Support

One Sample *t* test Usage?

Unfortunately...

- The one sample t test is essentially never used, but it servers a good purpose to familiarize us with the *t* distribution.
- Occasionally, we know both μ and σ, for instance with SAT or GRE scores, which necessitates the Z-test.
- Generally, we do not know μ or σ so the one sample *t* test is not frequently used.

Research and Statistical Support

One Sample *t* test Usage?

Unfortunately...

- The one sample t test is essentially never used, but it servers a good purpose to familiarize us with the *t* distribution.
- Occasionally, we know both μ and σ, for instance with SAT or GRE scores, which necessitates the Z-test.
- Generally, we do not know μ or σ so the one sample *t* test is not frequently used.
- Instead, as we will see; we estimate population values with sample statistics and compare samples to *infer* effects in the general population(s) of interest.

RSS
Research and Statistical Support

One Sample *t* test Usage?

Unfortunately...

- The one sample t test is essentially never used, but it servers a good purpose to familiarize us with the *t* distribution.
- Occasionally, we know both μ and σ, for instance with SAT or GRE scores, which necessitates the Z-test.
- Generally, we do not know μ or σ so the one sample *t* test is not frequently used.
- Instead, as we will see; we estimate population values with sample statistics and compare samples to *infer* effects in the general population(s) of interest.
    - The *t* distribution is used for other types of *t* tests which will be covered shortly.

RSS
Research and Statistical Support

Summary of Section 1

Section 1 covered the following topics:

- Introduced the *t* distribution

Summary of Section 1

Section 1 covered the following topics:

- Introduced the t distribution
- One Sample t test

Summary of Section 1

Section 1 covered the following topics:

- Introduced the t distribution
- One Sample t test
- Cohen's d Effect Size

RSS
Research and Statistical Support

Summary of Section 1

Section 1 covered the following topics:

- Introduced the *t* distribution
- One Sample *t* test
- Cohen's *d* Effect Size
- Use of delta (δ) for Statistical Power

Summary of Section 1

Section 1 covered the following topics:

- Introduced the t distribution
- One Sample t test
- Cohen's d Effect Size
- Use of delta (δ) for Statistical Power
- Use of delta (δ) for calculating a-priori sample size

RSS
Research and Statistical Support

Summary of Section 1

Section 1 covered the following topics:

- Introduced the *t* distribution
- One Sample *t* test
- Cohen's *d* Effect Size
- Use of delta (δ) for Statistical Power
- Use of delta (δ) for calculating a-priori sample size
- Calculation of Confidence Intervals.

RSS Research and Statistical Support

Dependent Samples *t* Test

- First, it goes by many names...

Dependent Samples *t* Test

- First, it goes by many names...
  - Dependent Samples *t* Test

RSS
Research and Statistical Support

Dependent Samples t Test

- First, it goes by many names...
  - Dependent Samples t Test
  - Matched-Pairs t Test

Dependent Samples *t* Test

- First, it goes by many names...
  - Dependent Samples *t* Test
  - Matched-Pairs *t* Test
  - Repeated Measures *t* Test

RSS
Research and Statistical Support

Dependent Samples *t* Test

- First, it goes by many names...
  - Dependent Samples *t* Test
  - Matched-Pairs *t* Test
  - Repeated Measures *t* Test
  - *t* Test for Dependent Means

Dependent Samples *t* Test

- First, it goes by many names...
  - Dependent Samples *t* Test
  - Matched-Pairs *t* Test
  - Repeated Measures *t* Test
  - *t* Test for Dependent Means
  - *t* Test for Related Samples

Dependent Samples *t* Test

- First, it goes by many names...
  - Dependent Samples *t* Test
  - Matched-Pairs *t* Test
  - Repeated Measures *t* Test
  - *t* Test for Dependent Means
  - *t* Test for Related Samples
- Essentially, it compares two sample means which are known to be related in some way.

RSS
Research and Statistical Support

Types of Related Samples

There are three types of related samples which are appropriate for the Dependent Samples *t* Test.

Types of Related Samples

There are three types of related samples which are appropriate for the Dependent Samples *t* Test.

- Natural Pairs: Comparing the scores of two groups of subjects which are related naturally.

Research and Statistical Support

Types of Related Samples

There are three types of related samples which are appropriate for the Dependent Samples *t* Test.

- Natural Pairs: Comparing the scores of two groups of subjects which are related naturally.
  - E.g., Twins (twin 1 vs. twin 2 on scores of trait anxiety).

RSS
Research and Statistical Support

Types of Related Samples

There are three types of related samples which are appropriate for the Dependent Samples *t* Test.

- Natural Pairs: Comparing the scores of two groups of subjects which are related naturally.
  - E.g., Twins (twin 1 vs. twin 2 on scores of trait anxiety).
- Matched Pairs: Pairs two groups of subjects on some characteristic that is likely to be important to the variable which you are studying.

RSS
Research and Statistical Support

Types of Related Samples

There are three types of related samples which are appropriate for the Dependent Samples *t* Test.

- Natural Pairs: Comparing the scores of two groups of subjects which are related naturally.
  - E.g., Twins (twin 1 vs. twin 2 on scores of trait anxiety).
- Matched Pairs: Pairs two groups of subjects on some characteristic that is likely to be important to the variable which you are studying.
  - E.g., Married couples (husbands vs. wives on ratings of marital satisfaction).

RSS
Research and Statistical Support

Types of Related Samples

There are three types of related samples which are appropriate for the Dependent Samples *t* Test.

- Natural Pairs: Comparing the scores of two groups of subjects which are related naturally.
  - E.g., Twins (twin 1 vs. twin 2 on scores of trait anxiety).
- Matched Pairs: Pairs two groups of subjects on some characteristic that is likely to be important to the variable which you are studying.
  - E.g., Married couples (husbands vs. wives on ratings of marital satisfaction).
- Repeated Measures: Same people at two different times of measure (of the same variable).

RSS
Research and Statistical Support

Types of Related Samples

There are three types of related samples which are appropriate for the Dependent Samples *t* Test.

- Natural Pairs: Comparing the scores of two groups of subjects which are related naturally.
  - E.g., Twins (twin 1 vs. twin 2 on scores of trait anxiety).
- Matched Pairs: Pairs two groups of subjects on some characteristic that is likely to be important to the variable which you are studying.
  - E.g., Married couples (husbands vs. wives on ratings of marital satisfaction).
- Repeated Measures: Same people at two different times of measure (of the same variable).
  - E.g., Pretest/post test (patients' life satisfaction ratings before and after treatment for depression).

RSS
Research and Statistical Support

Types of Related Samples

There are three types of related samples which are appropriate for the Dependent Samples *t* Test.

- Natural Pairs: Comparing the scores of two groups of subjects which are related naturally.
  - E.g., Twins (twin 1 vs. twin 2 on scores of trait anxiety).
- Matched Pairs: Pairs two groups of subjects on some characteristic that is likely to be important to the variable which you are studying.
  - E.g., Married couples (husbands vs. wives on ratings of marital satisfaction).
- Repeated Measures: Same people at two different times of measure (of the same variable).
  - E.g., Pretest/post test (patients' life satisfaction ratings before and after treatment for depression).

Gist: There is some sort of known meaningful relationship between the two groups of scores.

RSS Research and Statistical Support

Similiar to previous tests, but...

Similiar to previous tests, but...

- Now we are testing a **Difference Score.**

Similiar to previous tests, but...

- Now we are testing a **Difference Score**.
    - Instead of a mean; we are using the mean of differences.

Similiar to previous tests, but...

- Now we are testing a **Difference Score**.
  - Instead of a mean; we are using the mean of differences.
- When dealing with pairs of scores, we simply subtract one set of the pair from the other to get our *difference scores* for each pair.

RSS
Research and Statistical Support

Similiar to previous tests, but...

- Now we are testing a **Difference Score**.
  - Instead of a mean; we are using the mean of differences.
- When dealing with pairs of scores, we simply subtract one set of the pair from the other to get our *difference scores* for each pair.
  - $X - Y = D$ for the difference scores.

RSS
Research and Statistical Support

Similiar to previous tests, but...

- Now we are testing a **Difference Score**.
  - Instead of a mean; we are using the mean of differences.
- When dealing with pairs of scores, we simply subtract one set of the pair from the other to get our *difference scores* for each pair.
  - $X - Y = D$ for the difference scores.
  - It is the difference scores (D) with which we conduct the t test.

RSS
Research and Statistical Support

Similiar to previous tests, but...

- Now we are testing a **Difference Score**.
  - Instead of a mean; we are using the mean of differences.
- When dealing with pairs of scores, we simply subtract one set of the pair from the other to get our *difference scores* for each pair.
  - $X - Y = D$ for the difference scores.
  - It is the difference scores (D) with which we conduct the t test.
- The population of difference scores has $\mu = 0$ and we again estimate the standard deviation (σ).

Similiar to previous tests, but...

- Now we are testing a **Difference Score.**
  - Instead of a mean; we are using the mean of differences.
- When dealing with pairs of scores, we simply subtract one set of the pair from the other to get our *difference scores* for each pair.
  - $X - Y = D$ for the difference scores.
  - It is the difference scores (D) with which we conduct the t test.
- The population of difference scores has $\mu = 0$ and we again estimate the standard deviation (σ).
  - If there is no difference between the pairs, then the mean of the difference scores will be equal to zero.

RSS
Research and Statistical Support

Similiar to previous tests, but...

- Now we are testing a **Difference Score**.
  - Instead of a mean; we are using the mean of differences.
- When dealing with pairs of scores, we simply subtract one set of the pair from the other to get our *difference scores* for each pair.
  - $X - Y = D$ for the difference scores.
  - It is the difference scores (D) with which we conduct the t test.
- The population of difference scores has $\mu = 0$ and we again estimate the standard deviation (σ).
  - If there is no difference between the pairs, then the mean of the difference scores will be equal to zero.
  - Use the following notation:

RSS
Research and Statistical Support

Similiar to previous tests, but...

- Now we are testing a **Difference Score**.
  - Instead of a mean; we are using the mean of differences.
- When dealing with pairs of scores, we simply subtract one set of the pair from the other to get our *difference scores* for each pair.
  - $X - Y = D$ for the difference scores.
  - It is the difference scores (D) with which we conduct the t test.
- The population of difference scores has $\mu = 0$ and we again estimate the standard deviation (σ).
  - If there is no difference between the pairs, then the mean of the difference scores will be equal to zero.
  - Use the following notation:

$$\mu_D = 0 \quad \text{or} \quad \mu_1 - \mu_2 = 0$$

NHST Example

- We are interested in relationship satisfaction of young adults before and after they go off to college/university which separates them from their sweat-heart.

NHST Example

- We are interested in relationship satisfaction of young adults before and after they go off to college/university which separates them from their sweat-heart.
    - Relationship satisfaction rating: 0 - 50 range, higher score indicates greater satisfaction.

RSS
Research and Statistical Support

NHST Example

- We are interested in relationship satisfaction of young adults before and after they go off to college/university which separates them from their sweat-heart.
  - Relationship satisfaction rating: 0 - 50 range, higher score indicates greater satisfaction.
- Collect a sample of young adults which are about to be separated from their boy- or girl- friend by going to college.

NHST Example

- We are interested in relationship satisfaction of young adults before and after they go off to college/university which separates them from their sweat-heart.
  - Relationship satisfaction rating: 0 - 50 range, higher score indicates greater satisfaction.
- Collect a sample of young adults which are about to be separated from their boy- or girl- friend by going to college.
- Have them rate their satisfaction with the relationship.

NHST Example

- We are interested in relationship satisfaction of young adults before and after they go off to college/university which separates them from their sweat-heart.
  - Relationship satisfaction rating: 0 - 50 range, higher score indicates greater satisfaction.
- Collect a sample of young adults which are about to be separated from their boy- or girl- friend by going to college.
- Have them rate their satisfaction with the relationship.
- Then, after they have spent the first semester at colleges/universities (away from their sweat-heart), they again rate their relationship satisfaction.

RSS
Research and Statistical Support

Step 1: State the Null and Alternative Hypotheses

Define the populations: Relationship Satisfaction = RS

Step 1: State the Null and Alternative Hypotheses

Define the populations: Relationship Satisfaction = RS

- Population 1: RS prior to college separation.
- Population 2: RS after the first semester.

Step 1: State the Null and Alternative Hypotheses

Define the populations: Relationship Satisfaction = RS

- Population 1: RS prior to college separation.
- Population 2: RS after the first semester.

State the Hypotheses:

Step 1: State the Null and Alternative Hypotheses

Define the populations: Relationship Satisfaction = RS

- Population 1: RS prior to college separation.
- Population 2: RS after the first semester.

State the Hypotheses:

- The null hypothesis is: there will be no difference between the means of the ratings (before college vs. after the first semester).

Step 1: State the Null and Alternative Hypotheses

Define the populations: Relationship Satisfaction = RS

- Population 1: RS prior to college separation.
- Population 2: RS after the first semester.

State the Hypotheses:

- The null hypothesis is: there will be no difference between the means of the ratings (before college vs. after the first semester).

$$H_0 : \mu_1 - \mu_2 \leq 0 \qquad \text{or} \qquad \mu_D \leq 0$$

Step 1: State the Null and Alternative Hypotheses

Define the populations: Relationship Satisfaction = RS

- Population 1: RS prior to college separation.
- Population 2: RS after the first semester.

State the Hypotheses:

- The null hypothesis is: there will be no difference between the means of the ratings (before college vs. after the first semester).

$$H_0 : \mu_1 - \mu_2 \leq 0 \qquad \text{or} \qquad \mu_D \leq 0$$

- The alternative hypothesis is: there will be a significant **decrease** in the ratings.

RSS
Research and Statistical Support

Step 1: State the Null and Alternative Hypotheses

Define the populations: Relationship Satisfaction = RS

- Population 1: RS prior to college separation.
- Population 2: RS after the first semester.

State the Hypotheses:

- The null hypothesis is: there will be no difference between the means of the ratings (before college vs. after the first semester).

$$H_0 : \mu_1 - \mu_2 \leq 0 \qquad \text{or} \qquad \mu_D \leq 0$$

- The alternative hypothesis is: there will be a significant **decrease** in the ratings.

$$H_1 : \mu_1 - \mu_2 > 0 \qquad \text{or} \qquad \mu_D > 0$$

RSS
Research and Statistical Support

Step 1: State the Null and Alternative Hypotheses

Define the populations: Relationship Satisfaction = RS

- Population 1: RS prior to college separation.
- Population 2: RS after the first semester.

State the Hypotheses:

- The null hypothesis is: there will be no difference between the means of the ratings (before college vs. after the first semester).

$$H_0 : \mu_1 - \mu_2 \leq 0 \qquad \text{or} \qquad \mu_D \leq 0$$

- The alternative hypothesis is: there will be a significant **decrease** in the ratings.

$$H_1 : \mu_1 - \mu_2 > 0 \qquad \text{or} \qquad \mu_D > 0$$

Note the directional alternative hypothesis; pay careful attention to how we specified the hypotheses in symbols.

Step 1: State the Null and Alternative Hypotheses

Define the populations: Relationship Satisfaction = RS

- Population 1: RS prior to college separation.
- Population 2: RS after the first semester.

State the Hypotheses:

- The null hypothesis is: there will be no difference between the means of the ratings (before college vs. after the first semester).

$$H_0 : \mu_1 - \mu_2 \leq 0 \qquad \text{or} \qquad \mu_D \leq 0$$

- The alternative hypothesis is: there will be a significant **decrease** in the ratings.

$$H_1 : \mu_1 - \mu_2 > 0 \qquad \text{or} \qquad \mu_D > 0$$

Note the directional alternative hypothesis; pay careful attention to how we specified the hypotheses in symbols.

- If the alternative hypothesis expected an increase, then $H_1 : \mu_1 - \mu_2 < 0$ or $\mu_D < 0$

RSS
Research and Statistical Support

Step 2: Comparison Distribution (estimate σ with S_M).

Pair	Before	After	D	$\overline{D}$	$(D - \overline{D})$	$\left(D - \overline{D}\right)^2$
1	40	32	8	8	0	0
2	38	31	7	8	-1	1
3	36	30	6	8	-2	4
4	42	31	11	8	3	9
4			32			$SOS = 14$

RSS
Research and Statistical Support

Step 2: Comparison Distribution (estimate σ with S_M).

Pair	Before	After	D	$\overline{D}$	$(D - \overline{D})$	$\left(D - \overline{D}\right)^2$
1	40	32	8	8	0	0
2	38	31	7	8	-1	1
3	36	30	6	8	-2	4
4	42	31	11	8	3	9
4			32			$SOS = 14$

$$\overline{D} = \sum(D)/n_D = 32/4 = 8$$

RSS Research and Statistical Support

Step 2: Comparison Distribution (estimate σ with S_M).

Pair	Before	After	D	$\overline{D}$	$(D - \overline{D})$	$\left(D - \overline{D}\right)^2$
1	40	32	8	8	0	0
2	38	31	7	8	-1	1
3	36	30	6	8	-2	4
4	42	31	11	8	3	9
4			32			$SOS = 14$

$$\overline{D} = \sum(D)/n_D = 32/4 = 8$$

$$S^2 = SOS/df = 14/n_D - 1 = 14/3 = 4.67$$

RSS
Research and Statistical Support

Step 2: Comparison Distribution (estimate σ with S_M).

Pair	Before	After	D	$\overline{D}$	$(D - \overline{D})$	$\left(D - \overline{D}\right)^2$
1	40	32	8	8	0	0
2	38	31	7	8	-1	1
3	36	30	6	8	-2	4
4	42	31	11	8	3	9
4			32			$SOS = 14$

$$\overline{D} = \sum(D)/n_D = 32/4 = 8$$

$$S^2 = SOS/df = 14/n_D - 1 = 14/3 = 4.67$$

$$S_M^2 = \frac{S^2}{n_D} = \frac{4.67}{4} = 1.1675$$

RSS
Research and Statistical Support

Step 2: Comparison Distribution (estimate σ with S_M).

Pair	Before	After	D	$\overline{D}$	$(D-\overline{D})$	$\left(D-\overline{D}\right)^2$
1	40	32	8	8	0	0
2	38	31	7	8	-1	1
3	36	30	6	8	-2	4
4	42	31	11	8	3	9
4			32			$SOS = 14$

$$\overline{D} = \sum(D)/n_D = 32/4 = 8$$

$$S^2 = SOS/df = 14/n_D - 1 = 14/3 = 4.67$$

$$S_M^2 = \frac{S^2}{n_D} = \frac{4.67}{4} = 1.1675$$

$$S_M = \sqrt{S_M^2} = \sqrt{1.1675} = 1.08$$

RSS
Research and Statistical Support

Step 3: Determine the critical value

Step 3: Determine the critical value

- Deja-vu...?

Step 3: Determine the critical value

- Deja-vu...?
- Determine the critical t value.

Step 3: Determine the critical value

- Deja-vu...?
- Determine the critical t value.
- Significance level = .05, one-tailed test, $df = n_D - 1 = 3$.

RSS Research and Statistical Support

Step 3: Determine the critical value

- Deja-vu...?
- Determine the critical *t* value.
- Significance level = .05, one-tailed test, $df = n_D - 1 = 3$.
- Look in the appropriate column of the *t* table.

RSS
Research and Statistical Support

Step 3: Determine the critical value

- Deja-vu...?
- Determine the critical *t* value.
- Significance level = .05, one-tailed test, $df = n_D - 1 = 3$.
- Look in the appropriate column of the *t* table.
  - One-tailed, significance level = .05

    `http://www.math.unb.ca/~knight/utility/t-table.htm`

RSS
Research and Statistical Support

Step 3: Determine the critical value

- Deja-vu...?
- Determine the critical *t* value.
- Significance level = .05, one-tailed test, $df = n_D - 1 = 3$.
- Look in the appropriate column of the *t* table.
  - One-tailed, significance level = .05

    `http://www.math.unb.ca/~knight/utility/t-table.htm`
- Critical *t* value is 2.353

Research and Statistical Support

Step 3: Determine the critical value

- Deja-vu...?
- Determine the critical *t* value.
- Significance level = .05, one-tailed test, $df = n_D - 1 = 3$.
- Look in the appropriate column of the *t* table.
  - One-tailed, significance level = .05

    `http://www.math.unb.ca/~knight/utility/t-table.htm`
- Critical *t* value is 2.353
- Which means, we need t_{calc} to be greater than `|2.353|` (absolute value of 2.353) to find a significant difference.

Step 3: Determine the critical value

- Deja-vu...?
- Determine the critical *t* value.
- Significance level = .05, one-tailed test, $df = n_D - 1 = 3$.
- Look in the appropriate column of the *t* table.
  - One-tailed, significance level = .05

    `http://www.math.unb.ca/~knight/utility/t-table.htm`
- Critical *t* value is 2.353
- Which means, we need t_{calc} to be greater than $|2.353|$ (absolute value of 2.353) to find a significant difference.

Same critical value we used with the One Sample *t* Test.

RSS
Research and Statistical Support

Step 4: Calculate *t*

Step 4: Calculate *t*

- The general *t* formula:

RSS
Research and Statistical Support

Step 4: Calculate *t*

- The general *t* formula:

$$t = \frac{\overline{X} - \mu}{S_M}$$

RSS
Research and Statistical Support

Step 4: Calculate *t*

- The general *t* formula:

$$t = \frac{\overline{X} - \mu}{S_M}$$

- Remember we are dealing with the *difference scores (D)* but the *t* formula is essentially the same as above.

RSS
Research and Statistical Support

Step 4: Calculate *t*

- The general *t* formula:

$$t = \frac{\overline{X} - \mu}{S_M}$$

- Remember we are dealing with the *difference scores (D)* but the *t* formula is essentially the same as above.
- From previous calculations, we have:
  $\overline{D} = 8, \mu_D = 0, S_M = 1.08$

Step 4: Calculate t

- The general t formula:

$$t = \frac{\overline{X} - \mu}{S_M}$$

- Remember we are dealing with the *difference scores (D)* but the t formula is essentially the same as above.
- From previous calculations, we have: $\overline{D} = 8, \mu_D = 0, S_M = 1.08$

$$t = \frac{\overline{D} - \mu_D}{S_M} = \frac{8 - 0}{1.08} = 7.41$$

- So, our $t_{calc} = 7.41$ which is fairly large.

RSS
Research and Statistical Support

Step 5: Compare and make a decision.

- Since $t_{calc} = 7.41 > 2.353 = t_{crit}$ we reject the null hypothesis.

Step 5: Compare and make a decision.

- Since $t_{calc} = 7.41 > 2.353 = t_{crit}$ we reject the null hypothesis.
- The initial interpretation is...

Step 5: Compare and make a decision.

- Since $t_{calc} = 7.41 > 2.353 = t_{crit}$ we reject the null hypothesis.
- The initial interpretation is... There was a significant decrease in ratings of relationship satisfaction from before college separation ($M = 39$, $SD = 2.58$) to after the first semester of college separation ($M = 31$, $SD = 0.82$), $t(3) = 7.41, p = .005$ (one-tailed).

RSS
Research and Statistical Support

Step 5: Compare and make a decision.

- Since $t_{calc} = 7.41 > 2.353 = t_{crit}$ we reject the null hypothesis.
- The initial interpretation is... There was a significant decrease in ratings of relationship satisfaction from before college separation ($M = 39$, $SD = 2.58$) to after the first semester of college separation ($M = 31$, $SD = 0.82$), $t(3) = 7.41, p = .005$ (one-tailed).
  - One could also state the results as; $t(3) = 7.41, p < .05$ because .05 was our significance level.

RSS
Research and Statistical Support

Step 5: Compare and make a decision.

- Since $t_{calc} = 7.41 > 2.353 = t_{crit}$ we reject the null hypothesis.
- The initial interpretation is... There was a significant decrease in ratings of relationship satisfaction from before college separation ($M = 39$, $SD = 2.58$) to after the first semester of college separation ($M = 31$, $SD = 0.82$), $t(3) = 7.41, p = .005$ (one-tailed).
  - One could also state the results as; $t(3) = 7.41, p < .05$ because .05 was our significance level.
  - The 3 above is the degrees of freedom (*df*).

Step 5: Compare and make a decision.

- Since $t_{calc} = 7.41 > 2.353 = t_{crit}$ we reject the null hypothesis.
- The initial interpretation is... There was a significant decrease in ratings of relationship satisfaction from before college separation ($M = 39$, $SD = 2.58$) to after the first semester of college separation ($M = 31$, $SD = 0.82$), $t(3) = 7.41, p = .005$ (one-tailed).
  - One could also state the results as; $t(3) = 7.41, p < .05$ because .05 was our significance level.
  - The 3 above is the degrees of freedom (*df*).
- At this point in the course, you should be able to calculate the means and standard deviations of each group of scores; however the exact *p* value was obtained by verifying the results above in SPSS (see below).

RSS
Research and Statistical Support

Paired Samples Statistics

		Mean	N	Std. Deviation	Std. Error Mean
Pair 1	before	39.0000	4	2.58199	1.29099
	after	31.0000	4	.81650	.40825

Paired Samples Correlations

		N	Correlation	Sig.
Pair 1	before & after	4	.632	.368

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	before - after	8.00000	2.16025	1.08012	4.56257	11.43743	7.407	3	.005

RSS
Research and Statistical Support

Effect Size

- Recall, Cohen's *d*:

Effect Size

- Recall, Cohen's *d*:

$$d = \frac{\overline{X} - \mu}{\sigma}$$

Effect Size

- Recall, Cohen's d:

$$d = \frac{\overline{X} - \mu}{\sigma}$$

- Here, we do not know σ so we use S to estimate (not S_M because it is influenced heavily by sample size).

RSS
Research and Statistical Support

Effect Size

- Recall, Cohen's *d*:

$$d = \frac{\overline{X} - \mu}{\sigma}$$

- Here, we do not know σ so we use S to estimate (not S_M because it is influenced heavily by sample size).
    - Effect sizes are not influenced by sample size.

RSS
Research and Statistical Support

Effect Size

- Recall, Cohen's *d*:

$$d = \frac{\overline{X} - \mu}{\sigma}$$

- Here, we do not know σ so we use S to estimate (not S_M because it is influenced heavily by sample size).
  - Effect sizes are not influenced by sample size.
- Also, because we are dealing with difference scores, we use slightly different symbols in the formula:

Effect Size

- Recall, Cohen's *d*:

$$d = \frac{\overline{X} - \mu}{\sigma}$$

- Here, we do not know σ so we use S to estimate (not S_M because it is influenced heavily by sample size).
  - Effect sizes are not influenced by sample size.
- Also, because we are dealing with difference scores, we use slightly different symbols in the formula:

$$d = \frac{\overline{D} - \mu_D}{S}$$

Research and Statistical Support

Effect Size continued

- Cohen's *d* for difference scores:

Effect Size continued

- Cohen's *d* for difference scores:

$$d = \frac{\overline{D} - \mu_D}{S}$$

Effect Size continued

- Cohen's *d* for difference scores:

$$d = \frac{\overline{D} - \mu_D}{S}$$

- We have $S^2 = 4.67$ from above, which leads to an estimate of σ:

Effect Size continued

- Cohen's *d* for difference scores:

$$d = \frac{\overline{D} - \mu_D}{S}$$

- We have $S^2 = 4.67$ from above, which leads to an estimate of σ:

$$S = \sqrt{4.67} = 2.16$$

Effect Size continued

- Cohen's *d* for difference scores:

$$d = \frac{\overline{D} - \mu_D}{S}$$

- We have $S^2 = 4.67$ from above, which leads to an estimate of σ:

$$S = \sqrt{4.67} = 2.16$$

- Which in turn, leads to *d*:

RSS
Research and Statistical Support

Effect Size continued

- Cohen's d for difference scores:

$$d = \frac{\overline{D} - \mu_D}{S}$$

- We have $S^2 = 4.67$ from above, which leads to an estimate of σ:

$$S = \sqrt{4.67} = 2.16$$

- Which in turn, leads to d:

$$d = \frac{\overline{D} - \mu_D}{S} = \frac{8 - 0}{2.16} = 3.70$$

- A large effect size.

Using Delta (δ) for Statistical Power

- As was done in the one sample t test situation, here again we calculate δ as a combination of sample size and effect size and use it to look up the power in the δ table.

RSS
Research and Statistical Support

Using Delta (δ) for Statistical Power

- As was done in the one sample t test situation, here again we calculate δ as a combination of sample size and effect size and use it to look up the power in the δ table.
- However, because we now have **two** samples, the formula is slightly different.

RSS
Research and Statistical Support

Using Delta (δ) for Statistical Power

- As was done in the one sample *t* test situation, here again we calculate δ as a combination of sample size and effect size and use it to look up the power in the δ table.
- However, because we now have **two** samples, the formula is slightly different.

$$\delta = d * \sqrt{\frac{n}{2}}$$

RSS
Research and Statistical Support

Using Delta (δ) for Statistical Power

- As was done in the one sample *t* test situation, here again we calculate δ as a combination of sample size and effect size and use it to look up the power in the δ table.
- However, because we now have **two** samples, the formula is slightly different.

$$\delta = d * \sqrt{\frac{n}{2}}$$

- So, for our current example, we get:

Using Delta (δ) for Statistical Power

- As was done in the one sample *t* test situation, here again we calculate δ as a combination of sample size and effect size and use it to look up the power in the δ table.
- However, because we now have **two** samples, the formula is slightly different.

$$\delta = d * \sqrt{\frac{n}{2}}$$

- So, for our current example, we get:

$$\delta = d * \sqrt{\frac{n}{2}} = 3.7 * \sqrt{\frac{4}{2}} = 3.7 * \sqrt{2} = 3.7 * 1.414 = 5.233$$

RSS
Research and Statistical Support

Using Delta (δ) for Statistical Power

- As was done in the one sample *t* test situation, here again we calculate δ as a combination of sample size and effect size and use it to look up the power in the δ table.
- However, because we now have **two** samples, the formula is slightly different.

$$\delta = d * \sqrt{\frac{n}{2}}$$

- So, for our current example, we get:

$$\delta = d * \sqrt{\frac{n}{2}} = 3.7 * \sqrt{\frac{4}{2}} = 3.7 * \sqrt{2} = 3.7 * 1.414 = 5.233$$

- Since our δ table shows that anything with a $\delta > 5.00$ equates to power greater than 0.99, we can safely assume we have at least a power of 0.99.

Using Delta δ to calculate appropriate sample size

- The more useful way to use δ is for calculating adequate sample size during the planning of the study.

RSS
Research and Statistical Support

Using Delta δ to calculate appropriate sample size

- The more useful way to use δ is for calculating adequate sample size during the planning of the study.
  - First, we need to calculate δ for a desired power .60 with a one-tailed test at .05 significance level.

RSS
Research and Statistical Support

Using Delta δ to calculate appropriate sample size

- The more useful way to use δ is for calculating adequate sample size during the planning of the study.
  - First, we need to calculate δ for a desired power .60 with a one-tailed test at .05 significance level.

http://www.unt.edu/rss/class/Jon/ISSS_SC/Module008/delta

Using Delta δ to calculate appropriate sample size

- The more useful way to use δ is for calculating adequate sample size during the planning of the study.
  - First, we need to calculate δ for a desired power .60 with a one-tailed test at .05 significance level.

http://www.unt.edu/rss/class/Jon/ISSS_SC/Module008/delta

- So, $\delta = 1.90$, now we can calculate the sample size for a given effect size $d = .25$.

RSS
Research and Statistical Support

Using Delta δ to calculate appropriate sample size

- The more useful way to use δ is for calculating adequate sample size during the planning of the study.
  - First, we need to calculate δ for a desired power .60 with a one-tailed test at .05 significance level.

http://www.unt.edu/rss/class/Jon/ISSS_SC/Module008/delta

- So, $\delta = 1.90$, now we can calculate the sample size for a given effect size $d = .25$.

$$1.90 = .25 * \sqrt{n/2}$$

RSS Research and Statistical Support

Using Delta δ to calculate appropriate sample size

- The more useful way to use δ is for calculating adequate sample size during the planning of the study.
  - First, we need to calculate δ for a desired power .60 with a one-tailed test at .05 significance level.

http://www.unt.edu/rss/class/Jon/ISSS_SC/Module008/delta

- So, $\delta = 1.90$, now we can calculate the sample size for a given effect size $d = .25$.

$$1.90 = .25 * \sqrt{n/2}$$
$$1.90/.25 = \sqrt{n/2}$$

RSS Research and Statistical Support

Using Delta δ to calculate appropriate sample size

- The more useful way to use δ is for calculating adequate sample size during the planning of the study.
  - First, we need to calculate δ for a desired power .60 with a one-tailed test at .05 significance level.

http://www.unt.edu/rss/class/Jon/ISSS_SC/Module008/delta

- So, $\delta = 1.90$, now we can calculate the sample size for a given effect size $d = .25$.

$$1.90 = .25 * \sqrt{n/2}$$
$$1.90/.25 = \sqrt{n/2}$$
$$7.6^2 = n/2$$

RSS
Research and Statistical Support

Using Delta δ to calculate appropriate sample size

- The more useful way to use δ is for calculating adequate sample size during the planning of the study.
  - First, we need to calculate δ for a desired power .60 with a one-tailed test at .05 significance level.

http://www.unt.edu/rss/class/Jon/ISSS_SC/Module008/delta

- So, $\delta = 1.90$, now we can calculate the sample size for a given effect size $d = .25$.

$$1.90 = .25 * \sqrt{n/2}$$
$$1.90/.25 = \sqrt{n/2}$$
$$7.6^2 = n/2$$
$$2 * 57.76 = n$$

RSS Research and Statistical Support

Using Delta δ to calculate appropriate sample size

- The more useful way to use δ is for calculating adequate sample size during the planning of the study.
  - First, we need to calculate δ for a desired power .60 with a one-tailed test at .05 significance level.

http://www.unt.edu/rss/class/Jon/ISSS_SC/Module008/delta

- So, $\delta = 1.90$, now we can calculate the sample size for a given effect size $d = .25$.

$$1.90 = .25 * \sqrt{n/2}$$
$$1.90/.25 = \sqrt{n/2}$$
$$7.6^2 = n/2$$
$$2 * 57.76 = n$$
$$115.52 = n$$

RSS Research and Statistical Support

Calculating a Confidence Interval with $\overline{D}$

- Using essentially the same procedures we used with the one sample *t* test, we can calculate the lower limit (LL) and upper limit (UL).

Calculating a Confidence Interval with $\overline{D}$

- Using essentially the same procedures we used with the one sample *t* test, we can calculate the lower limit (LL) and upper limit (UL).
- Recall, the general formulas for a confidence interval are: LL = (-crit)*(SE) + mean and UL = (+crit)*(SE) + mean

RSS
Research and Statistical Support

Calculating a Confidence Interval with $\overline{D}$

- Using essentially the same procedures we used with the one sample *t* test, we can calculate the lower limit (LL) and upper limit (UL).
- Recall, the general formulas for a confidence interval are: LL = (-crit)*(SE) + mean and UL = (+crit)*(SE) + mean
- When in the Dependent Samples situation, we simply use the difference score mean.

RSS
Research and Statistical Support

Calculating a Confidence Interval with $\overline{D}$

- Using essentially the same procedures we used with the one sample *t* test, we can calculate the lower limit (LL) and upper limit (UL).
- Recall, the general formulas for a confidence interval are: LL = (-crit)*(SE) + mean and UL = (+crit)*(SE) + mean
- When in the Dependent Samples situation, we simply use the difference score mean.

$$LL = -t_{crit} * S_M + \overline{D} = -2.353 * 1.08 + 8 = -2.541 + 8 = 5.459$$

Calculating a Confidence Interval with $\overline{D}$

- Using essentially the same procedures we used with the one sample *t* test, we can calculate the lower limit (LL) and upper limit (UL).
- Recall, the general formulas for a confidence interval are: LL = (-crit)*(SE) + mean and UL = (+crit)*(SE) + mean
- When in the Dependent Samples situation, we simply use the difference score mean.

$$LL = -t_{crit} * S_M + \overline{D} = -2.353 * 1.08 + 8 = -2.541 + 8 = 5.459$$

$$UL = +t_{crit} * S_M + \overline{D} = +2.353 * 1.08 + 8 = +2.541 + 8 = 10.541$$

RSS
Research and Statistical Support

Interpretation of CI_95

- In this example, we calculated a 95% confidence interval (CI_{95}) because our critical value was based on a significance level of .05.

RSS
Research and Statistical Support

Interpretation of CI_95

- In this example, we calculated a 95% confidence interval (CI_{95}) because our critical value was based on a significance level of .05.

$$LL = -2.353 * 1.08 + 8 = 5.459$$
$$UL = +2.353 * 1.08 + 8 = 10.541$$

RSS
Research and Statistical Support

Interpretation of CI_95

- In this example, we calculated a 95% confidence interval (CI_{95}) because our critical value was based on a significance level of .05.

$$LL = -2.353 * 1.08 + 8 = 5.459$$
$$UL = +2.353 * 1.08 + 8 = 10.541$$

- If we drew an infinite number of samples of young adults' relationship satisfaction ratings, 95% of those samples' difference score means would be between 5.459 and 10.541.

RSS
Research and Statistical Support

Interpretation of CI_95

- In this example, we calculated a 95% confidence interval (CI_{95}) because our critical value was based on a significance level of .05.

$$LL = -2.353 * 1.08 + 8 = 5.459$$
$$UL = +2.353 * 1.08 + 8 = 10.541$$

- If we drew an infinite number of samples of young adults' relationship satisfaction ratings, 95% of those samples' difference score means would be between 5.459 and 10.541.
  - Remember, the population difference score mean is fixed (but unknown); while each sample has its own difference score mean (samples fluctuate).

RSS
Research and Statistical Support

Dependent Samples *t* Test Usage?

Fortunately...

Dependent Samples *t* Test Usage?

Fortunately...

- The Dependent samples t Test is quite powerful (i.e., statistical power).

RSS
Research and Statistical Support

Dependent Samples *t* Test Usage?

Fortunately...

- The Dependent samples t Test is quite powerful (i.e., statistical power).
- Repeated measures designs (or matched pairs, dependent samples, etc.) have more power due to less variance between the groups of scores.

RSS
Research and Statistical Support

Dependent Samples *t* Test Usage?

Fortunately...

- The Dependent samples t Test is quite powerful (i.e., statistical power).
- Repeated measures designs (or matched pairs, dependent samples, etc.) have more power due to less variance between the groups of scores.
  - The same individuals being tested at different times.

RSS
Research and Statistical Support

Dependent Samples *t* Test Usage?

Fortunately...

- The Dependent samples t Test is quite powerful (i.e., statistical power).
- Repeated measures designs (or matched pairs, dependent samples, etc.) have more power due to less variance between the groups of scores.
  - The same individuals being tested at different times.
  - But, there is no control group for comparison, so the results are somewhat limited.

RSS Research and Statistical Support

Dependent Samples *t* Test Usage?

Fortunately...

- The Dependent samples t Test is quite powerful (i.e., statistical power).
- Repeated measures designs (or matched pairs, dependent samples, etc.) have more power due to less variance between the groups of scores.
  - The same individuals being tested at different times.
  - But, there is no control group for comparison, so the results are somewhat limited.
- The *t* test in general is *robust* to errors.

Dependent Samples *t* Test Usage?

Fortunately...

- The Dependent samples t Test is quite powerful (i.e., statistical power).
- Repeated measures designs (or matched pairs, dependent samples, etc.) have more power due to less variance between the groups of scores.
  - The same individuals being tested at different times.
  - But, there is no control group for comparison, so the results are somewhat limited.
- The *t* test in general is *robust* to errors.
  - Robust (in this situation) means, even with moderate departures from normality we can be confident in our results.

RSS
Research and Statistical Support

Summary of Section 2

Section 2 covered the following topics:

RSS
Research and Statistical Support

Summary of Section 2

Section 2 covered the following topics:

- The Dependent Samples *t* Test.

Summary of Section 2

Section 2 covered the following topics:

- The Dependent Samples *t* Test.
- An NHST example of the Dependent Samples *t* Test.

Summary of Section 2

Section 2 covered the following topics:

- The Dependent Samples *t* Test.
- An NHST example of the Dependent Samples *t* Test.
- Cohen's *d* Effect Size

Summary of Section 2

Section 2 covered the following topics:

- The Dependent Samples *t* Test.
- An NHST example of the Dependent Samples *t* Test.
- Cohen's *d* Effect Size
- Use of delta (δ) for Statistical Power

RSS
Research and Statistical Support

Summary of Section 2

Section 2 covered the following topics:

- The Dependent Samples *t* Test.
- An NHST example of the Dependent Samples *t* Test.
- Cohen's *d* Effect Size
- Use of delta (δ) for Statistical Power
- Use of delta (δ) for calculating a-priori sample size

Summary of Section 2

Section 2 covered the following topics:

- The Dependent Samples t Test.
- An NHST example of the Dependent Samples t Test.
- Cohen's d Effect Size
- Use of delta (δ) for Statistical Power
- Use of delta (δ) for calculating a-priori sample size
- Calculation of Confidence Intervals.

Independent Samples t Test

- It too goes by many names...

RSS
Research and Statistical Support

Independent Samples t Test

- It too goes by many names...
  - The Independent Samples t Test
  - Independent Means t Test
  - Between Groups t Test
  - The t test for Independent Means...

RSS
Research and Statistical Support

Independent Samples t Test

- It too goes by many names...
  - The Independent Samples t Test
  - Independent Means t Test
  - Between Groups t Test
  - The t test for Independent Means...
- Essentially, it is used for comparing two sample means which are **not** related in some known or meaningful way.

RSS
Research and Statistical Support

Independent Samples *t* Test

- It too goes by many names...
  - The Independent Samples *t* Test
  - Independent Means *t* Test
  - Between Groups *t* Test
  - The *t* test for Independent Means...
- Essentially, it is used for comparing two sample means which are **not** related in some known or meaningful way.
  - Two Independent groups of scores.

Methodological Application

- The Independent Samples *t* Test is applicable when you have a dichotomous Independent Variable (IV) and an interval or ratio scaled Dependent Variable (DV).

RSS
Research and Statistical Support

Methodological Application

- The Independent Samples *t* Test is applicable when you have a dichotomous Independent Variable (IV) and an interval or ratio scaled Dependent Variable (DV).
  - One IV with two categories (sometimes called conditions).

Methodological Application

- The Independent Samples t Test is applicable when you have a dichotomous Independent Variable (IV) and an interval or ratio scaled Dependent Variable (DV).
  - One IV with two categories (sometimes called conditions).
  - One DV which is continuous or nearly continuous.

Methodological Application

- The Independent Samples t Test is applicable when you have a dichotomous Independent Variable (IV) and an interval or ratio scaled Dependent Variable (DV).
  - One IV with two categories (sometimes called conditions).
  - One DV which is continuous or nearly continuous.
- Evaluating two treatments for Schizophrenia.

RSS
Research and Statistical Support

Methodological Application

- The Independent Samples t Test is applicable when you have a dichotomous Independent Variable (IV) and an interval or ratio scaled Dependent Variable (DV).
  - One IV with two categories (sometimes called conditions).
  - One DV which is continuous or nearly continuous.
- Evaluating two treatments for Schizophrenia.
  - IV = Treatment (with two groups).

RSS
Research and Statistical Support

Methodological Application

- The Independent Samples *t* Test is applicable when you have a dichotomous Independent Variable (IV) and an interval or ratio scaled Dependent Variable (DV).
  - One IV with two categories (sometimes called conditions).
  - One DV which is continuous or nearly continuous.
- Evaluating two treatments for Schizophrenia.
  - IV = Treatment (with two groups).
    - Electro-Convulsive Therapy (ECT)

RSS
Research and Statistical Support

Methodological Application

- The Independent Samples *t* Test is applicable when you have a dichotomous Independent Variable (IV) and an interval or ratio scaled Dependent Variable (DV).
  - One IV with two categories (sometimes called conditions).
  - One DV which is continuous or nearly continuous.
- Evaluating two treatments for Schizophrenia.
  - IV = Treatment (with two groups).
    - Electro-Convulsive Therapy (ECT)
    - Insulin Shock Therapy (IST)

RSS
Research and Statistical Support

Methodological Application

- The Independent Samples *t* Test is applicable when you have a dichotomous Independent Variable (IV) and an interval or ratio scaled Dependent Variable (DV).
  - One IV with two categories (sometimes called conditions).
  - One DV which is continuous or nearly continuous.
- Evaluating two treatments for Schizophrenia.
  - IV = Treatment (with two groups).
    - Electro-Convulsive Therapy (ECT)
    - Insulin Shock Therapy (IST)
  - DV = Frequency of Hallucinations

RSS Research and Statistical Support

Application Distinctions

Application Distinctions

- **Dependent** Samples *t* Test: two groups of scores from the same people or people *related* in some meaningful, known way.

Application Distinctions

- **Dependent** Samples *t* Test: two groups of scores from the same people or people *related* in some meaningful, known way.
  - Same people at time 1 vs. time 2 or twin 1 vs. twin 2.

RSS
Research and Statistical Support

Application Distinctions

- **Dependent** Samples *t* Test: two groups of scores from the same people or people *related* in some meaningful, known way.
  - Same people at time 1 vs. time 2 or twin 1 vs. twin 2.
  - Comparison distribution: **Distribution of Difference Scores**.

RSS
Research and Statistical Support

Application Distinctions

- **Dependent** Samples *t* Test: two groups of scores from the same people or people *related* in some meaningful, known way.
  - Same people at time 1 vs. time 2 or twin 1 vs. twin 2.
  - Comparison distribution: **Distribution of Difference Scores**.
- **Independent** Samples *t* Test: two *independent* groups of people, each with a set of scores (i.e., group 1's scores vs. group 2's scores).

Application Distinctions

- **Dependent** Samples *t* Test: two groups of scores from the same people or people *related* in some meaningful, known way.
  - Same people at time 1 vs. time 2 or twin 1 vs. twin 2.
  - Comparison distribution: **Distribution of Difference Scores**.
- **Independent** Samples *t* Test: two *independent* groups of people, each with a set of scores (i.e., group 1's scores vs. group 2's scores).
  - A group of people exposed to one treatment vs. a group of people exposed to another treatment.

Application Distinctions

- **Dependent** Samples *t* Test: two groups of scores from the same people or people *related* in some meaningful, known way.
  - Same people at time 1 vs. time 2 or twin 1 vs. twin 2.
  - Comparison distribution: **Distribution of Difference Scores**.
- **Independent** Samples *t* Test: two *independent* groups of people, each with a set of scores (i.e., group 1's scores vs. group 2's scores).
  - A group of people exposed to one treatment vs. a group of people exposed to another treatment.
  - Comparison distribution: **Distribution of Differences Between Means**.

RSS
Research and Statistical Support

A Quick *note* about *notation*...

A Quick *note* about *notation*...

- With the Independent Samples *t* Test, we have two groups, identified with the notation:

A Quick *note* about *notation*...

- With the Independent Samples *t* Test, we have two groups, identified with the notation:
  - Group 1: n_1 $\overline{X}_1$ S_1^2 S_1 SOS_1 df_1

A Quick *note* about *notation*...

- With the Independent Samples t Test, we have two groups, identified with the notation:
  - Group 1: n_1 $\overline{X}_1$ S_1^2 S_1 SOS_1 df_1
  - Group 2: n_2 $\overline{X}_2$ S_2^2 S_2 SOS_2 df_2

RSS
Research and Statistical Support

A Quick *note* about *notation*...

- With the Independent Samples *t* Test, we have two groups, identified with the notation:
  - Group 1: n_1 $\overline{X}_1$ S_1^2 S_1 SOS_1 df_1
  - Group 2: n_2 $\overline{X}_2$ S_2^2 S_2 SOS_2 df_2
- So this: $S^2 = \frac{\sum(X-\overline{X})^2}{n-1}$ becomes: $S_1^2 = \frac{\sum(X_1-\overline{X}_1)^2}{n_1-1}$ for group 1.

A Quick *note* about *notation*...

- With the Independent Samples *t* Test, we have two groups, identified with the notation:
  - Group 1: n_1 $\overline{X}_1$ S_1^2 S_1 SOS_1 df_1
  - Group 2: n_2 $\overline{X}_2$ S_2^2 S_2 SOS_2 df_2
- So this: $S^2 = \frac{\sum(X-\overline{X})^2}{n-1}$ becomes: $S_1^2 = \frac{\sum(X_1-\overline{X_1})^2}{n_1-1}$ for group 1.
- Use subscripts to identify each group with either a 1 or a 2 subscript.

RSS
Research and Statistical Support

Getting to the Distribution of Differences Between Means (part 1).

Getting to the Distribution of Differences Between Means (part 1).

- Each group has a population distribution.

Getting to the Distribution of Differences Between Means (part 1).

- Each group has a population distribution.
  - We can estimate those populations' variances with the sample variances: S_1^2 and S_2^2

Research and Statistical Support

Getting to the Distribution of Differences Between Means (part 1).

- Each group has a population distribution.
  - We can estimate those populations' variances with the sample variances: S_1^2 and S_2^2
- Each group can be used to create a distribution of means.

RSS
Research and Statistical Support

Getting to the Distribution of Differences Between Means (part 1).

- Each group has a population distribution.
  - We can estimate those populations' variances with the sample variances: S_1^2 and S_2^2
- Each group can be used to create a distribution of means.
  - We can estimate the distribution of means' variances with the sample variances divided by the number of individuals in the samples S_{M1}^2 and S_{M1}^2

RSS
Research and Statistical Support

Getting to the Distribution of Differences Between Means (part 1).

- Each group has a population distribution.
  - We can estimate those populations' variances with the sample variances: S_1^2 and S_2^2
- Each group can be used to create a distribution of means.
  - We can estimate the distribution of means' variances with the sample variances divided by the number of individuals in the samples S_{M1}^2 and S_{M1}^2
- Using those two distributions of means, we can create a *Distribution of Differences Between Means*.

RSS
Research and Statistical Support

Getting to the Distribution of Differences Between Means (part 1).

- Each group has a population distribution.
  - We can estimate those populations' variances with the sample variances: S_1^2 and S_2^2
- Each group can be used to create a distribution of means.
  - We can estimate the distribution of means' variances with the sample variances divided by the number of individuals in the samples S_{M1}^2 and S_{M1}^2
- Using those two distributions of means, we can create a *Distribution of Differences Between Means*.
  BUT...

RSS
Research and Statistical Support

Getting to the Distribution of Differences Between Means (part 2).

Getting to the Distribution of Differences Between Means (part 2).

- Because we *assume* both σ_1^2 and σ_2^2 are equal; we must come up with an average of the two estimates S_1^2 and S_2^2 to get the best overall estimate of the population variance*.

RSS
Research and Statistical Support

Getting to the Distribution of Differences Between Means (part 2).

- Because we *assume* both σ_1^2 and σ_2^2 are equal; we must come up with an average of the two estimates S_1^2 and S_2^2 to get the best overall estimate of the population variance*.
  - ***This is especially crucial when the size of each group is different.**

RSS
Research and Statistical Support

Getting to the Distribution of Differences Between Means (part 2).

- Because we *assume* both σ_1^2 and σ_2^2 are equal; we must come up with an average of the two estimates S_1^2 and S_2^2 to get the best overall estimate of the population variance*.
  - ***This is especially crucial when the size of each group is different.**
- This *best estimate* is called the **pooled estimate** of the population variance.

RSS
Research and Statistical Support

Getting to the Distribution of Differences Between Means (part 2).

- Because we *assume* both σ_1^2 and σ_2^2 are equal; we must come up with an average of the two estimates S_1^2 and S_2^2 to get the best overall estimate of the population variance*.
  - ***This is especially crucial when the size of each group is different.**
- This *best estimate* is called the **pooled estimate** of the population variance.
  - Symbol: S_p^2

Research and Statistical Support

Getting to the Distribution of Differences Between Means (part 3).

Getting to the Distribution of Differences Between Means (part 3).

- To get S_p^2 you must get df_1, df_2, and df_{total} (also called df_t)

Getting to the Distribution of Differences Between Means (part 3).

- To get S_p^2 you must get df_1, df_2, and df_{total} (also called df_t)

$$df_1 = n_1 - 1 \qquad df_2 = n_2 - 1$$

Getting to the Distribution of Differences Between Means (part 3).

- To get S_p^2 you must get df_1, df_2, and df_{total} (also called df_t)

$$df_1 = n_1 - 1 \qquad df_2 = n_2 - 1$$

$$df_t = df_1 + df_2 \qquad \text{or} \qquad df_t = n_1 + n_2 - 2$$

Getting to the Distribution of Differences Between Means (part 3).

- To get S_p^2 you must get df_1, df_2, and df_{total} (also called df_t)

$$df_1 = n_1 - 1 \qquad df_2 = n_2 - 1$$

$$df_t = df_1 + df_2 \qquad \text{or} \qquad df_t = n_1 + n_2 - 2$$

- So, then we can get S_p^2 using:

Getting to the Distribution of Differences Between Means (part 3).

- To get S_p^2 you must get df_1, df_2, and df_{total} (also called df_t)

$$df_1 = n_1 - 1 \qquad df_2 = n_2 - 1$$

$$df_t = df_1 + df_2 \qquad \text{or} \qquad df_t = n_1 + n_2 - 2$$

- So, then we can get S_p^2 using:

$$S_p^2 = (S_1^2) * \frac{df_1}{df_t} + (S_2^2) * \frac{df_2}{df_t}$$

RSS
Research and Statistical Support

Getting to the Distribution of Differences Between Means (part 3).

- To get S_p^2 you must get df_1, df_2, and df_{total} (also called df_t)

$$df_1 = n_1 - 1 \qquad df_2 = n_2 - 1$$

$$df_t = df_1 + df_2 \qquad \text{or} \qquad df_t = n_1 + n_2 - 2$$

- So, then we can get S_p^2 using:

$$S_p^2 = \left(S_1^2\right) * \frac{df_1}{df_t} + \left(S_2^2\right) * \frac{df_2}{df_t}$$

- But wait...there's more...

RSS
Research and Statistical Support

Getting to the Distribution of Differences Between Means (part 4).

Getting to the Distribution of Differences Between Means (part 4).

- Now, we take S_p^2 and figure S_{M1}^2 and S_{M2}^2 by removing the influence of sample size.

RSS
Research and Statistical Support

Getting to the Distribution of Differences Between Means (part 4).

- Now, we take S_p^2 and figure S_{M1}^2 and S_{M2}^2 by removing the influence of sample size.

$$S_{M1}^2 = \frac{S_p^2}{n_1} \qquad S_{M2}^2 = \frac{S_p^2}{n_2}$$

Getting to the Distribution of Differences Between Means (part 4).

- Now, we take S_p^2 and figure S_{M1}^2 and S_{M2}^2 by removing the influence of sample size.

$$S_{M1}^2 = \frac{S_p^2}{n_1} \qquad S_{M2}^2 = \frac{S_p^2}{n_2}$$

- Which leads to...the variance of the distribution of differences between means.

RSS
Research and Statistical Support

Getting to the Distribution of Differences Between Means (part 4).

- Now, we take S_p^2 and figure S_{M1}^2 and S_{M2}^2 by removing the influence of sample size.

$$S_{M1}^2 = \frac{S_p^2}{n_1} \qquad S_{M2}^2 = \frac{S_p^2}{n_2}$$

- Which leads to...the variance of the distribution of differences between means.

$$S_{dif}^2 = S_{M1}^2 + S_{M2}^2$$

RSS
Research and Statistical Support

Getting to the Distribution of Differences Between Means (part 4).

- Now, we take S_p^2 and figure S_{M1}^2 and S_{M2}^2 by removing the influence of sample size.

$$S_{M1}^2 = \frac{S_p^2}{n_1} \qquad S_{M2}^2 = \frac{S_p^2}{n_2}$$

- Which leads to...the variance of the distribution of differences between means.

$$S_{dif}^2 = S_{M1}^2 + S_{M2}^2$$

- Which then leads to the standard deviation of the distribution of differences between means.

RSS
Research and Statistical Support

Getting to the Distribution of Differences Between Means (part 4).

- Now, we take S_p^2 and figure S_{M1}^2 and S_{M2}^2 by removing the influence of sample size.

$$S_{M1}^2 = \frac{S_p^2}{n_1} \qquad S_{M2}^2 = \frac{S_p^2}{n_2}$$

- Which leads to...the variance of the distribution of differences between means.

$$S_{dif}^2 = S_{M1}^2 + S_{M2}^2$$

- Which then leads to the standard deviation of the distribution of differences between means.

$$S_{dif} = \sqrt{S_{dif}^2}$$

RSS
Research and Statistical Support

Finally...the *t* test.

Finally...the *t* test.

- Now that we have S_{dif} we can calculate t or t_{calc}

Finally...the *t* test.

- Now that we have S_{dif} we can calculate t or t_{calc}

$$t_{calc} = \frac{\overline{X}_1 - \overline{X}_2}{S_{dif}}$$

RSS
Research and Statistical Support

Finally...the *t* test.

- Now that we have S_{dif} we can calculate t or t_{calc}

$$t_{calc} = \frac{\overline{X}_1 - \overline{X}_2}{S_{dif}}$$

- We would then use the df_t and our significance level to look in the t table to find our cutoff sample score (t_{crit})

RSS
Research and Statistical Support

Finally...the *t* test.

- Now that we have S_{dif} we can calculate t or t_{calc}

$$t_{calc} = \frac{\overline{X}_1 - \overline{X}_2}{S_{dif}}$$

- We would then use the df_t and our significance level to look in the t table to find our cutoff sample score (t_{crit})

`http://www.math.unb.ca/~knight/utility/t-table.htm`

RSS
Research and Statistical Support

NHST Example

NHST Example

- Examine whether viewers of John Stewart's *The Daily Show* know significantly **more** about world affairs than viewers of Bill O'Reilly's *The O'Reilly Factor* show.

RSS
Research and Statistical Support

NHST Example

- Examine whether viewers of John Stewart's *The Daily Show* know significantly **more** about world affairs than viewers of Bill O'Reilly's *The O'Reilly Factor* show.
- Randomly sample 16 cable viewers, randomly assign them to one of two *show* groups; Daily and Factor.

RSS
Research and Statistical Support

NHST Example

- Examine whether viewers of John Stewart's *The Daily Show* know significantly **more** about world affairs than viewers of Bill O'Reilly's *The O'Reilly Factor* show.
- Randomly sample 16 cable viewers, randomly assign them to one of two *show* groups; Daily and Factor.
- Have the participants watch 20 recent episodes of one show or the other, depending on their group assignment.

RSS
Research and Statistical Support

NHST Example

- Examine whether viewers of John Stewart's *The Daily Show* know significantly **more** about world affairs than viewers of Bill O'Reilly's *The O'Reilly Factor* show.
- Randomly sample 16 cable viewers, randomly assign them to one of two *show* groups; Daily and Factor.
- Have the participants watch 20 recent episodes of one show or the other, depending on their group assignment.
- Assess their knowledge of Current World Events using the CWE questionnaire, which has a range of 1 to 10.

Step 1

Step 1

- Define the populations and restate the research question as null and alternative hypotheses.

Step 1

- Define the populations and restate the research question as null and alternative hypotheses.
- Population 1: Americans who watch *The Daily Show*.

RSS
Research and Statistical Support

Step 1

- Define the populations and restate the research question as null and alternative hypotheses.
- Population 1: Americans who watch *The Daily Show*.
- Population 2: Americans who watch *The O'Reilly Factor*.

Research and Statistical Support

Step 1

- Define the populations and restate the research question as null and alternative hypotheses.
- Population 1: Americans who watch *The Daily Show*.
- Population 2: Americans who watch *The O'Reilly Factor*.

$$H_0 : \mu_1 \leq \mu_2 \text{ or } H_0 : \mu_1 - \mu_2 \leq 0$$

$$H_1 : \mu_1 > \mu_2 \text{ or } H_1 : \mu_1 - \mu_2 > 0$$

RSS
Research and Statistical Support

Step 1

- Define the populations and restate the research question as null and alternative hypotheses.
- Population 1: Americans who watch *The Daily Show*.
- Population 2: Americans who watch *The O'Reilly Factor*.

$$H_0 : \mu_1 \leq \mu_2 \text{ or } H_0 : \mu_1 - \mu_2 \leq 0$$

$$H_1 : \mu_1 > \mu_2 \text{ or } H_1 : \mu_1 - \mu_2 > 0$$

- In terms of knowledge about current events.

RSS
Research and Statistical Support

Step 1

- Define the populations and restate the research question as null and alternative hypotheses.
- Population 1: Americans who watch *The Daily Show*.
- Population 2: Americans who watch *The O'Reilly Factor*.

$$H_0 : \mu_1 \leq \mu_2 \text{ or } H_0 : \mu_1 - \mu_2 \leq 0$$

$$H_1 : \mu_1 > \mu_2 \text{ or } H_1 : \mu_1 - \mu_2 > 0$$

- In terms of knowledge about current events.
- Notice the directional alternative hypothesis (H_1) which indicates a one-tailed test.

RSS Research and Statistical Support

The Daily show group's data

Table 3: Daily Show Group

X_1	$\overline{X}_1$	$X_1 - \overline{X}_1$	$\left(X_1 - \overline{X}_1\right)^2$
6	6.75	-0.750	0.563
6	6.75	-0.750	0.563
9	6.75	2.250	5.063
8	6.75	1.250	1.563
4	6.75	-2.275	7.563
6	6.75	-0.750	0.563
7	6.75	0.250	0.063
8	6.75	1.250	1.563
$54 = \sum X_1$			$SOS_1 = 17.50$
$8 = n_1$			

$$S_1^2 = \frac{\sum\left(X_1 - \overline{X}_1\right)^2}{n_1 - 1} = \frac{SOS_1}{df_1} = \frac{17.50}{8-1} = \frac{17.50}{7} = 2.50$$

RSS
Research and Statistical Support

The Factor show group's data

Table 4: Factor Show Group

X_2	$\overline{X}_2$	$X_2 - \overline{X}_2$	$\left(X_2 - \overline{X}_2\right)^2$
5	3.875	1.125	1.266
4	3.875	0.125	0.016
3	3.875	-0.875	0.766
1	3.875	-2.875	8.266
5	3.875	1.125	1.266
6	3.875	2.125	4.516
3	3.875	-0.875	0.766
4	3.875	0.125	0.016
$31 = \sum X_2$			$SOS_2 = 16.875$
$8 = n_2$			

$$S_2^2 = \frac{\sum\left(X_2 - \overline{X}_2\right)^2}{n_2 - 1} = \frac{SOS_2}{df_2} = \frac{16.875}{8-1} = \frac{16.875}{7} = 2.411$$

RSS
Research and Statistical Support

Step 2(a)

- 2. Determine the characteristics of the comparison distribution.

RSS
Research and Statistical Support

Step 2(a)

- 2. Determine the characteristics of the comparison distribution.
- $df_t = df_1 + df_2 = n_1 + n_2 - 2 = 8 + 8 - 2 = 14$

RSS
Research and Statistical Support

Step 2(a)

- 2. Determine the characteristics of the comparison distribution.
- $df_t = df_1 + df_2 = n_1 + n_2 - 2 = 8 + 8 - 2 = 14$
- And from above, $S_1^2 = 2.500$ and $S_2^2 = 2.411$

RSS
Research and Statistical Support

Step 2(a)

- 2. Determine the characteristics of the comparison distribution.
- $df_t = df_1 + df_2 = n_1 + n_2 - 2 = 8 + 8 - 2 = 14$
- And from above, $S_1^2 = 2.500$ and $S_2^2 = 2.411$
- (a) Calculate the pooled estimate of the population variance.

RSS
Research and Statistical Support

Step 2(a)

- 2. Determine the characteristics of the comparison distribution.
- $df_t = df_1 + df_2 = n_1 + n_2 - 2 = 8 + 8 - 2 = 14$
- And from above, $S_1^2 = 2.500$ and $S_2^2 = 2.411$
- (a) Calculate the pooled estimate of the population variance.

$$S_p^2 = (S_1^2) * \frac{df_1}{df_t} + (S_2^2) * \frac{df_2}{df_t} = (2.500) * \frac{7}{14} + (2.411) * \frac{7}{14} = 2.4555$$

Step 2(a)

- 2. Determine the characteristics of the comparison distribution.
- $df_t = df_1 + df_2 = n_1 + n_2 - 2 = 8 + 8 - 2 = 14$
- And from above, $S_1^2 = 2.500$ and $S_2^2 = 2.411$
- (a) Calculate the pooled estimate of the population variance.

$$S_p^2 = (S_1^2) * \frac{df_1}{df_t} + (S_2^2) * \frac{df_2}{df_t} = (2.500) * \frac{7}{14} + (2.411) * \frac{7}{14} = 2.4555$$

- So, $S_p^2 = 2.46$

RSS
Research and Statistical Support

Step 2(b)

- (b) Calculate the variance of each distribution of means:

RSS
Research and Statistical Support

Step 2(b)

- (b) Calculate the variance of each distribution of means:

$$S^2_{M1} = \frac{S^2_p}{n_1} = \frac{2.46}{8} = 0.3075$$

$$S^2_{M2} = \frac{S^2_p}{n_2} = \frac{2.46}{8} = 0.3075$$

RSS
Research and Statistical Support

Step 2(b)

- (b) Calculate the variance of each distribution of means:

$$S_{M1}^2 = \frac{S_p^2}{n_1} = \frac{2.46}{8} = 0.3075$$

$$S_{M2}^2 = \frac{S_p^2}{n_2} = \frac{2.46}{8} = 0.3075$$

- Please note; if the groups were different sizes, the variances of each distribution of means would be different.

RSS
Research and Statistical Support

Step 2(c) and Step 2(d)

- (c) Calculate the variance of the distribution of differences between means:

Step 2(c) and Step 2(d)

- (c) Calculate the variance of the distribution of differences between means:

$$S^2_{dif} = S^2_{M1} + S^2_{M2} = 0.3075 + 0.3075 = 0.615$$

RSS
Research and Statistical Support

Step 2(c) and Step 2(d)

- (c) Calculate the variance of the distribution of differences between means:

$$S^2_{dif} = S^2_{M1} + S^2_{M2} = 0.3075 + 0.3075 = 0.615$$

- So, $S^2_{dif} = 0.62$

RSS
Research and Statistical Support

Step 2(c) and Step 2(d)

- (c) Calculate the variance of the distribution of differences between means:

$$S^2_{dif} = S^2_{M1} + S^2_{M2} = 0.3075 + 0.3075 = 0.615$$

- So, $S^2_{dif} = 0.62$
- (d) Calculate the standard deviation of the distribution of differences between means:

RSS
Research and Statistical Support

Step 2(c) and Step 2(d)

- (c) Calculate the variance of the distribution of differences between means:

$$S^2_{dif} = S^2_{M1} + S^2_{M2} = 0.3075 + 0.3075 = 0.615$$

- So, $S^2_{dif} = 0.62$
- (d) Calculate the standard deviation of the distribution of differences between means:

$$S_{dif} = \sqrt{S^2_{dif}} = \sqrt{.62} = 0.78$$

RSS
Research and Statistical Support

Step 2(c) and Step 2(d)

- (c) Calculate the variance of the distribution of differences between means:

$$S^2_{dif} = S^2_{M1} + S^2_{M2} = 0.3075 + 0.3075 = 0.615$$

- So, $S^2_{dif} = 0.62$
- (d) Calculate the standard deviation of the distribution of differences between means:

$$S_{dif} = \sqrt{S^2_{dif}} = \sqrt{.62} = 0.78$$

- So, $S_{dif} = 0.78$

Step 3

- 3. Determine the critical sample score on the comparison distribution at which the null hypothesis should be rejected.

RSS
Research and Statistical Support

Step 3

- 3. Determine the critical sample score on the comparison distribution at which the null hypothesis should be rejected.
  - Significance level = .05

RSS
Research and Statistical Support

Step 3

- 3. Determine the critical sample score on the comparison distribution at which the null hypothesis should be rejected.
  - Significance level = .05
  - Two-tailed test (based on H_1).

RSS
Research and Statistical Support

Step 3

- 3. Determine the critical sample score on the comparison distribution at which the null hypothesis should be rejected.

  - Significance level = .05
  - Two-tailed test (based on H_1).
  - $df_t = n_1 + n_2 - 2 = 7 + 7 - 2 = 14$

RSS
Research and Statistical Support

Step 3

- 3. Determine the critical sample score on the comparison distribution at which the null hypothesis should be rejected.
  - Significance level = .05
  - Two-tailed test (based on H_1).
  - $df_t = n_1 + n_2 - 2 = 7 + 7 - 2 = 14$

http://www.math.unb.ca/~knight/utility/t-table.htm

RSS
Research and Statistical Support

Step 3

- 3. Determine the critical sample score on the comparison distribution at which the null hypothesis should be rejected.
  - Significance level = .05
  - Two-tailed test (based on H_1).
  - $df_t = n_1 + n_2 - 2 = 7 + 7 - 2 = 14$

http://www.math.unb.ca/~knight/utility/t-table.htm

- $t_{crit} = 1.761$

RSS
Research and Statistical Support

Step 4

- 4. Determine the sample's score on the comparison distribution:

Step 4

- 4. Determine the sample's score on the comparison distribution:
  - Compute t_{calc}

Step 4

- 4. Determine the sample's score on the comparison distribution:
  - Compute t_{calc}

$$t = \frac{\overline{X}_1 - \overline{X}_2}{S_{dif}} = \frac{6.75 - 3.875}{0.78} = 3.69$$

Step 4

- 4. Determine the sample's score on the comparison distribution:
  - Compute t_{calc}

$$t = \frac{\overline{X}_1 - \overline{X}_2}{S_{dif}} = \frac{6.75 - 3.875}{0.78} = 3.69$$

- So, $t_{calc} = 3.69$

RSS Research and Statistical Support

Step 5

- 5. Compare the scores from Step 3 and Step 4, and make a decision to reject the null hypothesis or fail to reject the null hypothesis.

RSS
Research and Statistical Support

Step 5

- 5. Compare the scores from Step 3 and Step 4, and make a decision to reject the null hypothesis or fail to reject the null hypothesis.
- Because; $t_{calc} = 3.69 > 1.761 = t_{crit}$ we reject the null hypothesis and conclude there was a statistically significant difference between the two show groups.

RSS
Research and Statistical Support

Step 5

- 5. Compare the scores from Step 3 and Step 4, and make a decision to reject the null hypothesis or fail to reject the null hypothesis.
- Because; $t_{calc} = 3.69 > 1.761 = t_{crit}$ we reject the null hypothesis and conclude there was a statistically significant difference between the two show groups.
  - But, you should know by now, that's not the whole story.

RSS
Research and Statistical Support

Calculating Effect Size for two Independent Groups

- Recall, the general formula for Cohen's *d*.

Calculating Effect Size for two Independent Groups

- Recall, the general formula for Cohen's *d*.

$$d = \frac{\mu_1 - \mu_2}{\sigma}$$

RSS
Research and Statistical Support

Calculating Effect Size for two Independent Groups

- Recall, the general formula for Cohen's *d*.

$$d = \frac{\mu_1 - \mu_2}{\sigma}$$

- In the current (independent groups) situation, we have:

Calculating Effect Size for two Independent Groups

- Recall, the general formula for Cohen's *d*.

$$d = \frac{\mu_1 - \mu_2}{\sigma}$$

- In the current (independent groups) situation, we have:

$$d = \frac{\overline{X}_1 - \overline{X}_2}{S_p} = \frac{6.75 - 3.875}{\sqrt{2.46}} = \frac{2.875}{1.57} = 1.83$$

RSS
Research and Statistical Support

Calculating Effect Size for two Independent Groups

- Recall, the general formula for Cohen's *d*.

$$d = \frac{\mu_1 - \mu_2}{\sigma}$$

- In the current (independent groups) situation, we have:

$$d = \frac{\overline{X}_1 - \overline{X}_2}{S_p} = \frac{6.75 - 3.875}{\sqrt{2.46}} = \frac{2.875}{1.57} = 1.83$$

- So, the effect size is fairly large; $d = 1.83$

Using Delta (δ) for Statistical Power

- As was done with the Dependent Samples *t* Test situation, here again we calculate δ as a combination of sample size and effect size and use it to look up the power in the δ table.

RSS
Research and Statistical Support

Using Delta (δ) for Statistical Power

- As was done with the Dependent Samples *t* Test situation, here again we calculate δ as a combination of sample size and effect size and use it to look up the power in the δ table.
- We use the exact same formula as was used with Dependent Samples (Section 2 above).

RSS
Research and Statistical Support

Using Delta (δ) for Statistical Power

- As was done with the Dependent Samples *t* Test situation, here again we calculate δ as a combination of sample size and effect size and use it to look up the power in the δ table.
- We use the exact same formula as was used with Dependent Samples (Section 2 above).

$$\delta = d * \sqrt{\frac{n}{2}}$$

Research and Statistical Support

Using Delta (δ) for Statistical Power

- As was done with the Dependent Samples *t* Test situation, here again we calculate δ as a combination of sample size and effect size and use it to look up the power in the δ table.
- We use the exact same formula as was used with Dependent Samples (Section 2 above).

$$\delta = d * \sqrt{\frac{n}{2}}$$

- So, for our current example, we get the following (where *n* is the number **per group**):

Using Delta (δ) for Statistical Power

- As was done with the Dependent Samples *t* Test situation, here again we calculate δ as a combination of sample size and effect size and use it to look up the power in the δ table.
- We use the exact same formula as was used with Dependent Samples (Section 2 above).

$$\delta = d * \sqrt{\frac{n}{2}}$$

- So, for our current example, we get the following (where *n* is the number **per group**):

$$\delta = d * \sqrt{\frac{n}{2}} = 1.83 * \sqrt{\frac{8}{2}} = 1.83 * \sqrt{4} = 1.83 * 2 = 3.66$$

RSS
Research and Statistical Support

Using Delta (δ) for Statistical Power

- As was done with the Dependent Samples *t* Test situation, here again we calculate δ as a combination of sample size and effect size and use it to look up the power in the δ table.
- We use the exact same formula as was used with Dependent Samples (Section 2 above).

$$\delta = d * \sqrt{\frac{n}{2}}$$

- So, for our current example, we get the following (where *n* is the number **per group**):

$$\delta = d * \sqrt{\frac{n}{2}} = 1.83 * \sqrt{\frac{8}{2}} = 1.83 * \sqrt{4} = 1.83 * 2 = 3.66$$

- The δ table shows that with a $\delta = 3.60$ (note: it is best to round down) we have a power of 0.98.

RSS
Research and Statistical Support

http://www.unt.edu/rss/class/Jon/ISSS_SC/Module008/delta

Using Delta δ to calculate appropriate sample size

- The more useful way to use δ is for calculating adequate sample size during the planning of the study.

RSS
Research and Statistical Support

Using Delta δ to calculate appropriate sample size

- The more useful way to use δ is for calculating adequate sample size during the planning of the study.
  - First, we need to calculate δ for a desired power .60 with a one-tailed test at .05 significance level.

RSS
Research and Statistical Support

Using Delta δ to calculate appropriate sample size

- The more useful way to use δ is for calculating adequate sample size during the planning of the study.
  - First, we need to calculate δ for a desired power .60 with a one-tailed test at .05 significance level.

http://www.unt.edu/rss/class/Jon/ISSS_SC/Module008/delta

Using Delta δ to calculate appropriate sample size

- The more useful way to use δ is for calculating adequate sample size during the planning of the study.
  - First, we need to calculate δ for a desired power .60 with a one-tailed test at .05 significance level.

`http://www.unt.edu/rss/class/Jon/ISSS_SC/Module008/delta`

- So, $\delta = 1.90$, now we can calculate the sample size for a given effect size $d = .25$.

Using Delta δ to calculate appropriate sample size

- The more useful way to use δ is for calculating adequate sample size during the planning of the study.
  - First, we need to calculate δ for a desired power .60 with a one-tailed test at .05 significance level.

http://www.unt.edu/rss/class/Jon/ISSS_SC/Module008/delta

- So, $\delta = 1.90$, now we can calculate the sample size for a given effect size $d = .25$.

$$1.90 = .25 * \sqrt{n/2}$$

RSS Research and Statistical Support

Using Delta δ to calculate appropriate sample size

- The more useful way to use δ is for calculating adequate sample size during the planning of the study.
  - First, we need to calculate δ for a desired power .60 with a one-tailed test at .05 significance level.

http://www.unt.edu/rss/class/Jon/ISSS_SC/Module008/delta

- So, $\delta = 1.90$, now we can calculate the sample size for a given effect size $d = .25$.

$$1.90 = .25 * \sqrt{n/2}$$
$$1.90/.25 = \sqrt{n/2}$$

RSS
Research and Statistical Support

Using Delta δ to calculate appropriate sample size

- The more useful way to use δ is for calculating adequate sample size during the planning of the study.
  - First, we need to calculate δ for a desired power .60 with a one-tailed test at .05 significance level.

http://www.unt.edu/rss/class/Jon/ISSS_SC/Module008/delta

- So, $\delta = 1.90$, now we can calculate the sample size for a given effect size $d = .25$.

$$1.90 = .25 * \sqrt{n/2}$$
$$1.90/.25 = \sqrt{n/2}$$
$$7.6^2 = n/2$$

RSS
Research and Statistical Support

Using Delta δ to calculate appropriate sample size

- The more useful way to use δ is for calculating adequate sample size during the planning of the study.
  - First, we need to calculate δ for a desired power .60 with a one-tailed test at .05 significance level.

http://www.unt.edu/rss/class/Jon/ISSS_SC/Module008/delta

- So, $\delta = 1.90$, now we can calculate the sample size for a given effect size $d = .25$.

$$\begin{aligned} 1.90 &= .25 * \sqrt{n/2} \\ 1.90/.25 &= \sqrt{n/2} \\ 7.6^2 &= n/2 \\ 2 * 57.76 &= n \end{aligned}$$

RSS Research and Statistical Support

Using Delta δ to calculate appropriate sample size

- The more useful way to use δ is for calculating adequate sample size during the planning of the study.
  - First, we need to calculate δ for a desired power .60 with a one-tailed test at .05 significance level.

http://www.unt.edu/rss/class/Jon/ISSS_SC/Module008/delta

- So, $\delta = 1.90$, now we can calculate the sample size for a given effect size $d = .25$.

$$1.90 = .25 * \sqrt{n/2}$$
$$1.90/.25 = \sqrt{n/2}$$
$$7.6^2 = n/2$$
$$2 * 57.76 = n$$
$$115.52 = n$$

RSS Research and Statistical Support

Using Delta δ to calculate appropriate sample size

- The more useful way to use δ is for calculating adequate sample size during the planning of the study.
  - First, we need to calculate δ for a desired power .60 with a one-tailed test at .05 significance level.

http://www.unt.edu/rss/class/Jon/ISSS_SC/Module008/delta

- So, $\delta = 1.90$, now we can calculate the sample size for a given effect size $d = .25$.

$$1.90 = .25 * \sqrt{n/2}$$
$$1.90/.25 = \sqrt{n/2}$$
$$7.6^2 = n/2$$
$$2 * 57.76 = n$$
$$115.52 = n$$

- As you can see, this is exactly as we did for the dependent samples situation. Just remember that the n refers to the number of **each group**.

An Additional comments on Power

- *t* tests with evenly distributed participants have greater power than those where the participants are unevenly distributed.

RSS
Research and Statistical Support

An Additional comments on Power

- *t* tests with evenly distributed participants have greater power than those where the participants are unevenly distributed.
- The dependent samples design has greater power (all else being equal, such as sample size, effect size, etc.) than the independent samples design.

RSS
Research and Statistical Support

An Additional comments on Power

- *t* tests with evenly distributed participants have greater power than those where the participants are unevenly distributed.
- The dependent samples design has greater power (all else being equal, such as sample size, effect size, etc.) than the independent samples design.
- And, as always, the larger the sample size, the greater the power.

Calculating a Confidence Interval

- Note, we are calculating the interval on the difference between means.

Calculating a Confidence Interval

- Note, we are calculating the interval on the difference between means.
- Recall there are two parts of a confidence interval, the upper limit (UL) and the lower limit (LL).

Calculating a Confidence Interval

- Note, we are calculating the interval on the difference between means.
- Recall there are two parts of a confidence interval, the upper limit (UL) and the lower limit (LL).
- The general form of the equations for each limit are:

RSS
Research and Statistical Support

Calculating a Confidence Interval

- Note, we are calculating the interval on the difference between means.
- Recall there are two parts of a confidence interval, the upper limit (UL) and the lower limit (LL).
- The general form of the equations for each limit are:

$$LL = (-crit) * (SE) + mean$$
$$UL = (+crit) * (SE) + mean$$

RSS
Research and Statistical Support

Calculating a Confidence Interval

- Note, we are calculating the interval on the difference between means.
- Recall there are two parts of a confidence interval, the upper limit (UL) and the lower limit (LL).
- The general form of the equations for each limit are:

$$LL = (-crit) * (SE) + mean$$
$$UL = (+crit) * (SE) + mean$$

- In the current situation for the differences between means:

RSS
Research and Statistical Support

Calculating a Confidence Interval

- Note, we are calculating the interval on the difference between means.
- Recall there are two parts of a confidence interval, the upper limit (UL) and the lower limit (LL).
- The general form of the equations for each limit are:

$$LL = (-crit) * (SE) + mean$$
$$UL = (+crit) * (SE) + mean$$

- In the current situation for the differences between means:

$$LL = (-t_{crit}) * (S_{dif}) + \left(\overline{X}_1 - \overline{X}_2\right) = -1.761 * 0.78 + (6.75 - 3.875) = -1.374 + 2.875 = 1.501$$
$$UL = (+t_{crit}) * (S_{dif}) + \left(\overline{X}_1 - \overline{X}_2\right) = +1.761 * 0.78 + (6.75 - 3.875) = +1.374 + 2.875 = 4.249$$
$$LL = 1.50$$
$$UL = 4.25$$

RSS
Research and Statistical Support

Interpretation of Confidence Interval

- Recall, we had a significance level of .05 (t_{crit}), so we conducted a 95% confidence interval (CI_{95}) on the differences between means.

Interpretation of Confidence Interval

- Recall, we had a significance level of .05 (t_{crit}), so we conducted a 95% confidence interval (CI_{95}) on the differences between means.
- The Lower Limit was 1.50 and the Upper Limit was 4.25.

RSS
Research and Statistical Support

Interpretation of Confidence Interval

- Recall, we had a significance level of .05 (t_{crit}), so we conducted a 95% confidence interval (CI_{95}) on the differences between means.
- The Lower Limit was 1.50 and the Upper Limit was 4.25.
- So, if we drew an infinite number of random samples of viewers of each show, 95% of the differences between means would be between 1.50 and 4.25.

RSS
Research and Statistical Support

Interpretation of Confidence Interval

- Recall, we had a significance level of .05 (t_{crit}), so we conducted a 95% confidence interval (CI_{95}) on the differences between means.
- The Lower Limit was 1.50 and the Upper Limit was 4.25.
- So, if we drew an infinite number of random samples of viewers of each show, 95% of the differences between means would be between 1.50 and 4.25.
  - Remember, the mean of the population of differences between means is fixed (but unknown); while each sample has its own differences between means (samples fluctuate).

RSS
Research and Statistical Support

Summary of Section 3

Section 3 covered the following topics:

Summary of Section 3

Section 3 covered the following topics:

- The Independent Samples t Test.

RSS
Research and Statistical Support

Summary of Section 3

Section 3 covered the following topics:

- The Independent Samples *t* Test.
- An NHST example of the Independent Samples *t* Test.

Summary of Section 3

Section 3 covered the following topics:

- The Independent Samples t Test.
- An NHST example of the Independent Samples t Test.
- Cohen's d Effect Size

Summary of Section 3

Section 3 covered the following topics:

- The Independent Samples *t* Test.
- An NHST example of the Independent Samples *t* Test.
- Cohen's *d* Effect Size
- Use of delta (δ) for Statistical Power

RSS
Research and Statistical Support

Summary of Section 3

Section 3 covered the following topics:

- The Independent Samples *t* Test.
- An NHST example of the Independent Samples *t* Test.
- Cohen's *d* Effect Size
- Use of delta (δ) for Statistical Power
- Use of delta (δ) for calculating a-priori sample size

Summary of Section 3

Section 3 covered the following topics:

- The Independent Samples *t* Test.
- An NHST example of the Independent Samples *t* Test.
- Cohen's *d* Effect Size
- Use of delta (δ) for Statistical Power
- Use of delta (δ) for calculating a-priori sample size
- Calculation of Confidence Intervals.

Assumptions for *t* tests

All statistical analyses have assumptions. If the assumptions are violated, then the analysis is invalidated; meaning, we must either choose a different analysis or modify the existing one in order to have faith in the results.

RSS
Research and Statistical Support

Assumptions for *t* tests

All statistical analyses have assumptions. If the assumptions are violated, then the analysis is invalidated; meaning, we must either choose a different analysis or modify the existing one in order to have faith in the results.

- Generally speaking, all three *t* tests have the same assumptions:

RSS
Research and Statistical Support

Assumptions for *t* tests

All statistical analyses have assumptions. If the assumptions are violated, then the analysis is invalidated; meaning, we must either choose a different analysis or modify the existing one in order to have faith in the results.

- Generally speaking, all three *t* tests have the same assumptions:
    - Normality

RSS
Research and Statistical Support

Assumptions for *t* tests

All statistical analyses have assumptions. If the assumptions are violated, then the analysis is invalidated; meaning, we must either choose a different analysis or modify the existing one in order to have faith in the results.

- Generally speaking, all three *t* tests have the same assumptions:
  - Normality
  - Homogeneity of Variance

RSS
Research and Statistical Support

Assumptions for *t* tests

All statistical analyses have assumptions. If the assumptions are violated, then the analysis is invalidated; meaning, we must either choose a different analysis or modify the existing one in order to have faith in the results.

- Generally speaking, all three *t* tests have the same assumptions:
  - Normality
  - Homogeneity of Variance
  - Independence

RSS
Research and Statistical Support

Normality assumption

- The normality assumption refers to the population(s) distribution(s), which are assumed to be normal.

Normality assumption

- The normality assumption refers to the population(s) distribution(s), which are assumed to be normal.
- In the one sample situation, we assume the population distribution is normal.

Normality assumption

- The normality assumption refers to the population(s) distribution(s), which are assumed to be normal.
- In the one sample situation, we assume the population distribution is normal.
- In the dependent samples situation, we assume the distribution of differences is normally distributed.

RSS
Research and Statistical Support

Normality assumption

- The normality assumption refers to the population(s) distribution(s), which are assumed to be normal.
- In the one sample situation, we assume the population distribution is normal.
- In the dependent samples situation, we assume the distribution of differences is normally distributed.
- In the independent samples situation, we assume the distribution of differences between means is normal.

Normality assumption

- The normality assumption refers to the population(s) distribution(s), which are assumed to be normal.
- In the one sample situation, we assume the population distribution is normal.
- In the dependent samples situation, we assume the distribution of differences is normally distributed.
- In the independent samples situation, we assume the distribution of differences between means is normal.
- The t tests can tolerate some deviation from normality, but if skewness and/or kurtosis are greater than |1.0| in the sample(s) then you should be cautious about proceeding.

RSS Research and Statistical Support

Homogeneity of Variances (HOV) assumption

- Does not apply to the one sample *t* test, but does with the dependent and independent tests.

Research and Statistical Support

Homogeneity of Variances (HOV) assumption

- Does not apply to the one sample *t* test, but does with the dependent and independent tests.
- This assumption states that each of the two populations (represented by our two sample groups) have equal (homogeneous) variances.

RSS
Research and Statistical Support

Homogeneity of Variances (HOV) assumption

- Does not apply to the one sample *t* test, but does with the dependent and independent tests.
- This assumption states that each of the two populations (represented by our two sample groups) have equal (homogeneous) variances.
  - As the definition implies, the variance of one group should be very similar to the variance of the other group.

Homogeneity of Variances (HOV) assumption

- Does not apply to the one sample *t* test, but does with the dependent and independent tests.
- This assumption states that each of the two populations (represented by our two sample groups) have equal (homogeneous) variances.
  - As the definition implies, the variance of one group should be very similar to the variance of the other group.
  - There are empirical tests for this assumption in virtually all statistical software packages.

RSS
Research and Statistical Support

Homogeneity of Variances (HOV) assumption

- Does not apply to the one sample *t* test, but does with the dependent and independent tests.
- This assumption states that each of the two populations (represented by our two sample groups) have equal (homogeneous) variances.
    - As the definition implies, the variance of one group should be very similar to the variance of the other group.
    - There are empirical tests for this assumption in virtually all statistical software packages.
- If the assumption is violated, there are options available.

RSS
Research and Statistical Support

Homogeneity of Variances (HOV) assumption

- Does not apply to the one sample *t* test, but does with the dependent and independent tests.
- This assumption states that each of the two populations (represented by our two sample groups) have equal (homogeneous) variances.
  - As the definition implies, the variance of one group should be very similar to the variance of the other group.
  - There are empirical tests for this assumption in virtually all statistical software packages.
- If the assumption is violated, there are options available.
  - Use trimmed samples; trimming 10 or 20% of the extreme scores (from each group) to make them more homogeneous.

RSS
Research and Statistical Support

Homogeneity of Variances (HOV) assumption

- Does not apply to the one sample *t* test, but does with the dependent and independent tests.
- This assumption states that each of the two populations (represented by our two sample groups) have equal (homogeneous) variances.
    - As the definition implies, the variance of one group should be very similar to the variance of the other group.
    - There are empirical tests for this assumption in virtually all statistical software packages.
- If the assumption is violated, there are options available.
    - Use trimmed samples; trimming 10 or 20% of the extreme scores (from each group) to make them more homogeneous.
    - Use a corrected *t* test formula, such as *t* prime (t'):

RSS
Research and Statistical Support

Homogeneity of Variances (HOV) assumption

- Does not apply to the one sample *t* test, but does with the dependent and independent tests.
- This assumption states that each of the two populations (represented by our two sample groups) have equal (homogeneous) variances.
  - As the definition implies, the variance of one group should be very similar to the variance of the other group.
  - There are empirical tests for this assumption in virtually all statistical software packages.
- If the assumption is violated, there are options available.
  - Use trimmed samples; trimming 10 or 20% of the extreme scores (from each group) to make them more homogeneous.
  - Use a corrected *t* test formula, such as *t* prime (t'):

$$t' = \frac{\overline{X}_1 - \overline{X}_2}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}}$$

RSS
Research and Statistical Support

Independence assumption

- This assumption refers to the independence of observations (i.e. scores) in the population(s); meaning, each score should not be related to any other score in the group (or sample in the one sample situation).

Independence assumption

- This assumption refers to the independence of observations (i.e. scores) in the population(s); meaning, each score should not be related to any other score in the group (or sample in the one sample situation).
- Typically, this is handled by random sampling (and random assignment to groups in the case of the independent samples situation).

RSS
Research and Statistical Support

Independence assumption

- This assumption refers to the independence of observations (i.e. scores) in the population(s); meaning, each score should not be related to any other score in the group (or sample in the one sample situation).
- Typically, this is handled by random sampling (and random assignment to groups in the case of the independent samples situation).
- If your study did not use random sampling, then you may have a problem with biased results.

RSS
Research and Statistical Support

Independence assumption

- This assumption refers to the independence of observations (i.e. scores) in the population(s); meaning, each score should not be related to any other score in the group (or sample in the one sample situation).
- Typically, this is handled by random sampling (and random assignment to groups in the case of the independent samples situation).
- If your study did not use random sampling, then you may have a problem with biased results.
    - For instance, if you offer compensation for participation (i.e. paying people to be in your study) and an entire community of underprivileged folks show up, then your results are not likely to be applicable to a larger population (i.e. more privileged folks).

RSS
Research and Statistical Support

Multiple Tests in one Study

- Family-wise error rate refers to the amount of Type I error (α) associated with multiple tests in one study.

Multiple Tests in one Study

- Family-wise error rate refers to the amount of Type I error (α) associated with multiple tests in one study.
  - Symbol for a Family-wise error rate of 0.05: $FW_{.05}$

RSS
Research and Statistical Support

Multiple Tests in one Study

- Family-wise error rate refers to the amount of Type I error (α) associated with multiple tests in one study.
  - Symbol for a Family-wise error rate of 0.05: $FW_{.05}$
- Family-wise error rate can be controlled by planning for it prior to data collection.

RSS
Research and Statistical Support

Multiple Tests in one Study

- Family-wise error rate refers to the amount of Type I error (α) associated with multiple tests in one study.
  - Symbol for a Family-wise error rate of 0.05: $FW_{.05}$
- Family-wise error rate can be controlled by planning for it prior to data collection.
- Say you want to conduct 3 *t* tests in one study and you want to use an $\alpha = .05$.

RSS
Research and Statistical Support

Multiple Tests in one Study

- Family-wise error rate refers to the amount of Type I error (α) associated with multiple tests in one study.
  - Symbol for a Family-wise error rate of 0.05: $FW_{.05}$
- Family-wise error rate can be controlled by planning for it prior to data collection.
- Say you want to conduct 3 *t* tests in one study and you want to use an $\alpha = .05$.
- Then, $FW_{.05} = .05/3 = 0.0167$

RSS
Research and Statistical Support

Multiple Tests in one Study

- Family-wise error rate refers to the amount of Type I error (α) associated with multiple tests in one study.
  - Symbol for a Family-wise error rate of 0.05: $FW_{.05}$
- Family-wise error rate can be controlled by planning for it prior to data collection.
- Say you want to conduct 3 *t* tests in one study and you want to use an $\alpha = .05$.
- Then, $FW_{.05} = .05/3 = 0.0167$
- The 0.0167 is called the *pair-wise* error rate.

RSS
Research and Statistical Support

Multiple Tests in one Study

- Family-wise error rate refers to the amount of Type I error (α) associated with multiple tests in one study.
  - Symbol for a Family-wise error rate of 0.05: $FW_{.05}$
- Family-wise error rate can be controlled by planning for it prior to data collection.
- Say you want to conduct 3 *t* tests in one study and you want to use an $\alpha = .05$.
- Then, $FW_{.05} = .05/3 = 0.0167$
- The 0.0167 is called the *pair-wise* error rate.
  - Each *t* test is done with $\alpha = 0.0167$ so total error rate is limited to .05 (i.e. $0.0167 * 3 = .05$).

RSS
Research and Statistical Support

Multiple Tests in one Study

- Family-wise error rate refers to the amount of Type I error (α) associated with multiple tests in one study.
    - Symbol for a Family-wise error rate of 0.05: $FW_{.05}$
- Family-wise error rate can be controlled by planning for it prior to data collection.
- Say you want to conduct 3 *t* tests in one study and you want to use an $\alpha = .05$.
- Then, $FW_{.05} = .05/3 = 0.0167$
- The 0.0167 is called the *pair-wise* error rate.
    - Each *t* test is done with $\alpha = 0.0167$ so total error rate is limited to .05 (i.e. $0.0167 * 3 = .05$).
    - However, such strict significance levels make rejecting the null difficult.

RSS Research and Statistical Support

More *t* Tests

- The discussion above (i.e. Family-wise & Pair-wise error rates) is important for two reasons.

RSS
Research and Statistical Support

More *t* Tests

- The discussion above (i.e. Family-wise & Pair-wise error rates) is important for two reasons.
  - First, as mentioned on the previous slide, you may want to conduct multiple *t* tests in one study; as your primary analyses.

RSS
Research and Statistical Support

More *t* Tests

- The discussion above (i.e. Family-wise & Pair-wise error rates) is important for two reasons.
  - First, as mentioned on the previous slide, you may want to conduct multiple *t* tests in one study; as your primary analyses.
    - The three types of *t* tests discussed so far are the *primary t* tests; meaning, they can be used as the primary analysis of a study.

More *t* Tests

- The discussion above (i.e. Family-wise & Pair-wise error rates) is important for two reasons.
  - First, as mentioned on the previous slide, you may want to conduct multiple *t* tests in one study; as your primary analyses.
    - The three types of *t* tests discussed so far are the *primary t* tests; meaning, they can be used as the primary analysis of a study.
  - However, in future modules you will see there are many more forms of *t* tests which are used as *secondary* analyses; meaning, they support a primary analysis.

RSS
Research and Statistical Support

More *t* Tests

- The discussion above (i.e. Family-wise & Pair-wise error rates) is important for two reasons.
  - First, as mentioned on the previous slide, you may want to conduct multiple *t* tests in one study; as your primary analyses.
    - The three types of *t* tests discussed so far are the *primary t* tests; meaning, they can be used as the primary analysis of a study.
  - However, in future modules you will see there are many more forms of *t* tests which are used as *secondary* analyses; meaning, they support a primary analysis.
    - In the ANOVA situation *t* tests can be used for discovering pair-wise differences among multiple groups.

RSS
Research and Statistical Support

More *t* Tests

- The discussion above (i.e. Family-wise & Pair-wise error rates) is important for two reasons.
  - First, as mentioned on the previous slide, you may want to conduct multiple *t* tests in one study; as your primary analyses.
    - The three types of *t* tests discussed so far are the *primary t* tests; meaning, they can be used as the primary analysis of a study.
  - However, in future modules you will see there are many more forms of *t* tests which are used as *secondary* analyses; meaning, they support a primary analysis.
    - In the ANOVA situation *t* tests can be used for discovering pair-wise differences among multiple groups.
    - In the Regression situation *t* tests are used to evaluate the significance of the weights associated with predictors in a model.

More *t* Tests

- The discussion above (i.e. Family-wise & Pair-wise error rates) is important for two reasons.
  - First, as mentioned on the previous slide, you may want to conduct multiple *t* tests in one study; as your primary analyses.
    - The three types of *t* tests discussed so far are the *primary t* tests; meaning, they can be used as the primary analysis of a study.
  - However, in future modules you will see there are many more forms of *t* tests which are used as *secondary* analyses; meaning, they support a primary analysis.
    - In the ANOVA situation *t* tests can be used for discovering pair-wise differences among multiple groups.
    - In the Regression situation *t* tests are used to evaluate the significance of the weights associated with predictors in a model.
- The bottom line is this, *t* tests are important and are frequently used; and the issue of cumulative error rates is important.

RSS
Research and Statistical Support

Summary of Section 4

Section 4 covered the following topics:

Summary of Section 4

Section 4 covered the following topics:

- The Assumptions of t Tests

Summary of Section 4

Section 4 covered the following topics:

- The Assumptions of *t* Tests
- Family-wise and Pair-wise Error rates

RSS Research and Statistical Support

Summary of Module 8

Module 8 covered the following topics:

- One Sample *t* Test

Summary of Module 8

Module 8 covered the following topics:

- One Sample t Test
- Dependent Samples t Test

RSS
Research and Statistical Support

Summary of Module 8

Module 8 covered the following topics:

- One Sample *t* Test
- Dependent Samples *t* Test
- Independent Samples *t* Test

Summary of Module 8

Module 8 covered the following topics:

- One Sample *t* Test
- Dependent Samples *t* Test
- Independent Samples *t* Test
- Summary of *t* Tests

Summary of Module 8

Module 8 covered the following topics:

- One Sample *t* Test
- Dependent Samples *t* Test
- Independent Samples *t* Test
- Summary of *t* Tests

A firm understanding of the topics covered here and previously will be necessary for understanding future topics.

This concludes Module 8

Next time Module 9.

- Next time we'll begin covering Introduction to Analysis of Variance.
- Until next time; have a nice day.

These slides initially created on: October 12, 2010
These slides last updated on: October 15, 2010

- The bottom date shown is the date this Adobe.pdf file was created; LaTeX[1] has a command for automatically inserting the date of a document's creation.

[1] This document was created in LaTeX using the Beamer package