Introduction to Probability and Statistics Eleventh Edition

Chapter 10
Inference from Small Samples

Some graphic screen captures from *Seeing Statistics* ®

Introduction

- When the sample size is small, the estimation and testing procedures of Chapters 8 and 9 are not appropriate.
- There are equivalent small sample test and estimation procedures for
  - ✓ μ, the mean of a normal population
  - ✓ $\mu_1 - \mu_2$, the difference between two population means
    - Case 1: 2 independent random samples
    - Case 2: 2 dependent random samples (a paired-difference test)

The Sampling Distribution of the Sample Mean

- If n is large, the sampling distribution of $\bar{x}$ is approximately normal. The test statistic

$$\boxed{\frac{\bar{x}-\mu}{s/\sqrt{n}}}$$

is **approximately normally distributed.**

- But if n is not large, then this test statistic **is not approximately normal**.

$$\boxed{\frac{\bar{x}-\mu}{s/\sqrt{n}} \text{ is not normal!}}$$

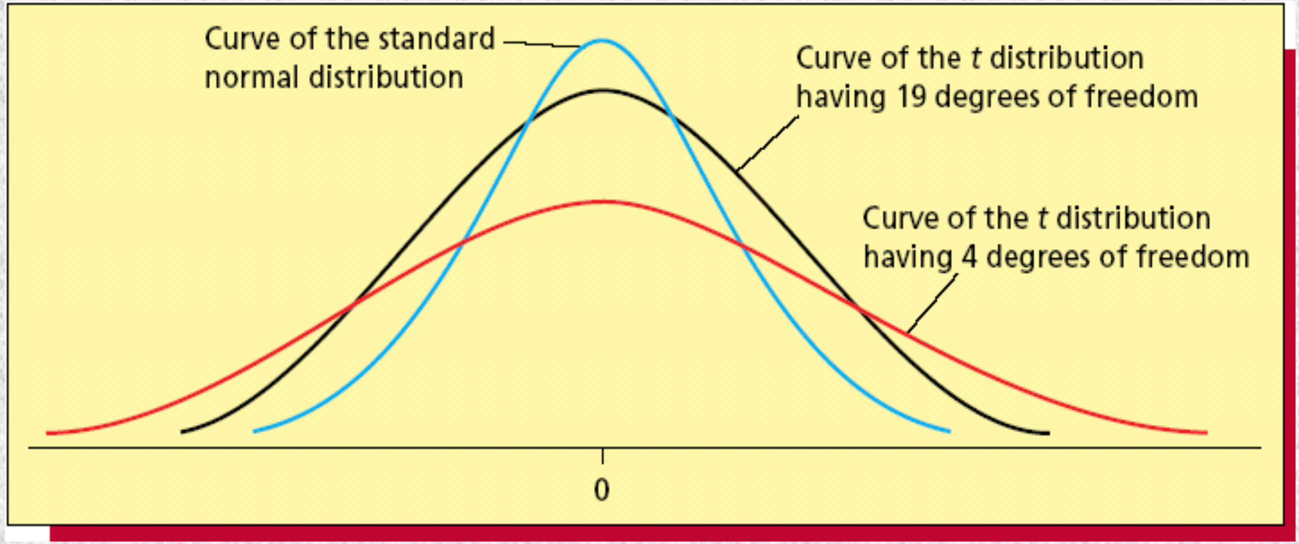
Student's *t* Distribution

- Fortunately, this statistic does have a sampling distribution that is well known to statisticians, called the **Student's t distribution,** with ***n*-1 degrees of freedom.**

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

- We can use this distribution to create estimation and testing procedures for the population mean μ.

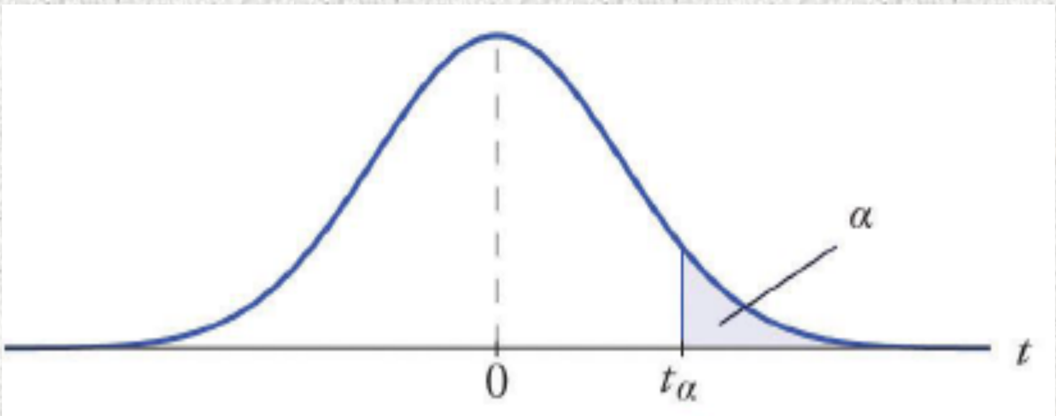
Properties of Student's *t*



- **Mound-shaped** and symmetric about 0.
- **More variable than *z***, with "heavier tails"
- Shape depends on the sample size *n* or the **degrees of freedom (*df.*), *n*-1.**
- As *n* increases the shapes of the *t* and *z* distributions become almost identical.

Using the *t*-Table

- The table gives the values of t that cut off certain critical values in the tail of the t distribution.
- Index ***df*** and the appropriate tail area ***a*** to find t_a, the value of t with area ***a*** to its right.



Using the *t*-Table

df	$t_{.100}$	$t_{.050}$	$t_{.025}$	$t_{.010}$
1	3.078	6.314	12.706	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
7	1.415	1.895	2.365	2.998
8	1.397	1.860	2.306	2.896
9	1.383	1.833	2.262	2.821
10	1.372	1.812	2.228	2.764
11	1.363	1.796	2.201	2.718
12	1.356	1.782	2.179	2.681
13	1.350	1.771	2.160	2.650
14	1.345	1.761	2.145	2.624
15	1.341	1.753	2.131	2.602

For a random sample of size n = 10, find a value of t that cuts off .025 in the right tail.

Row = $df = n - 1 = 9$

Column subscript = a = .025

$t_{.025} = 2.262$

Small Sample Inference for a Population Mean μ

- The basic procedures are the same as those used for large samples.
- Assume that a random sample is taken from a normal population.
- For a test of hypothesis:

Test $\mathrm{H}_0 : \mu = \mu_0$ versus H_a : one or two tailed

using the test statistic

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

using p - values or a rejection region based on

a t - distribution with $df = n - 1$.

Small Sample Inference for a Population Mean μ

- For a 100(1−α)% confidence interval for the population mean μ:

$$\bar{x} \pm t_{\alpha/2} \frac{s}{\sqrt{n}}$$

where $t_{\alpha/2}$ is the value of t that cuts off area $\alpha/2$ in the tail of a t-distribution with $df = n - 1$.

Example 1

A sprinkler system is designed so that the average time for the sprinklers to activate after being turned on is no more than 15 seconds. A test of 6 systems gave the following times:

17, 31, 12, 17, 13, 25 seconds

Is the system not working as specified? Test using $\alpha = 0.05$. Assume that the time is normally distributed.

$H_0: \mu = 15$ (working as specified)
$H_a: \mu > 15$ (not working as specified)

Example 1

Data: 17, 31, 12, 17, 13, 25

First, calculate the sample mean and standard deviation, using your calculator or the formulas in Chapter 2.

$$\bar{x} = \frac{\sum x_i}{n} = \frac{115}{6} = 19.167$$

$$s = \sqrt{\frac{\sum x_i^2 - \frac{\left(\sum x\right)^2}{n}}{n-1}} = \sqrt{\frac{2477 - \frac{115^2}{6}}{5}} = 7.387$$

Example 1

Data: 17, 31, 12, 17, 13, 25

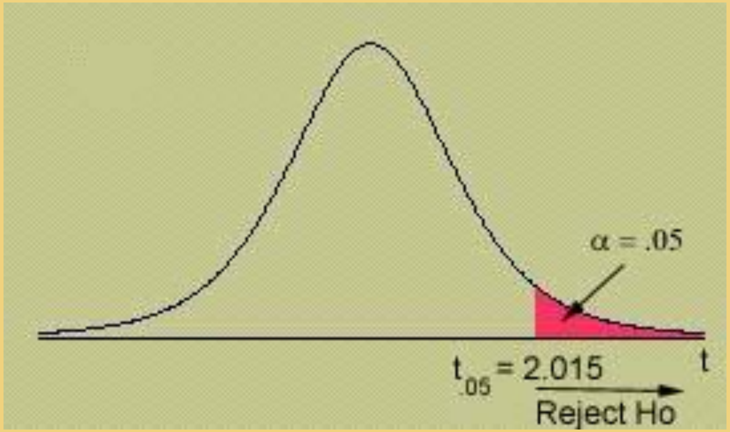
Calculate the test statistic and find the rejection region for $\alpha = 0.05$.

Test Statistic :

$$t = \frac{\bar{x} - \mu_0}{s/\sqrt{n}} = \frac{19.167 - 15}{7.387/\sqrt{6}} = 1.38$$

Degrees of freedom:

$$df = n - 1 = 5$$



Rejection Region:

Reject H_0 if $t > 2.015$.

Conclusion

Data: 17, 31, 12, 17, 13, 25

Compare the observed test statistic to the rejection region, and draw conclusions.

$$H_0 : \mu = 15$$
$$H_a : \mu > 15$$

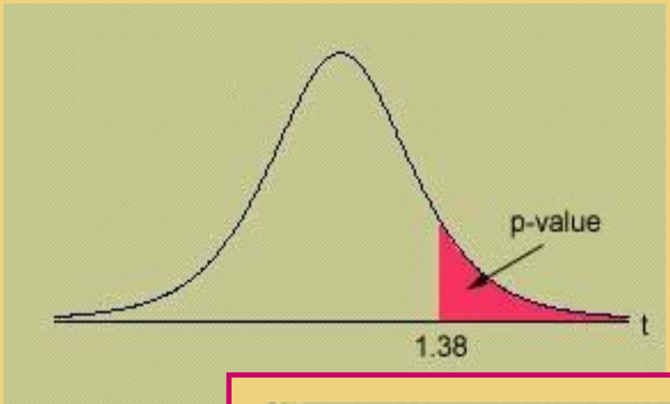
Test statistic : $t = 1.38$

Rejection Region: Reject H_0 if $t > 2.015$

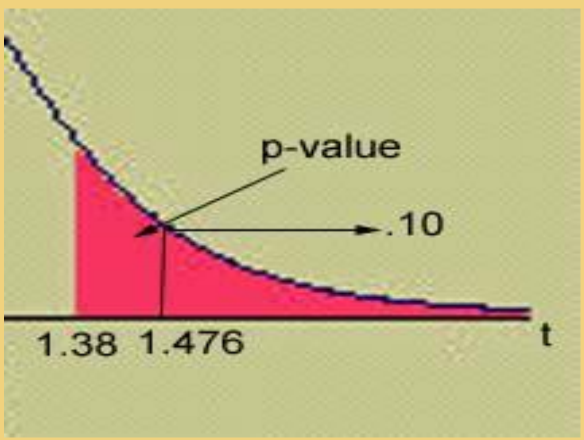
Conclusion: For our example, $t = 1.38$ does not fall in the rejection region and H_0 is not rejected. There is insufficient evidence to indicate that the average activation time is greater than 15 seconds.

Approximating the *p*-value

- You can only approximate the *p*-value for the test using *t*-table.



df	$t_{.100}$	$t_{.050}$
1	3.078	6.314
2	1.886	2.920
3	1.638	2.353
4	1.533	2.132
5	1.476	2.015



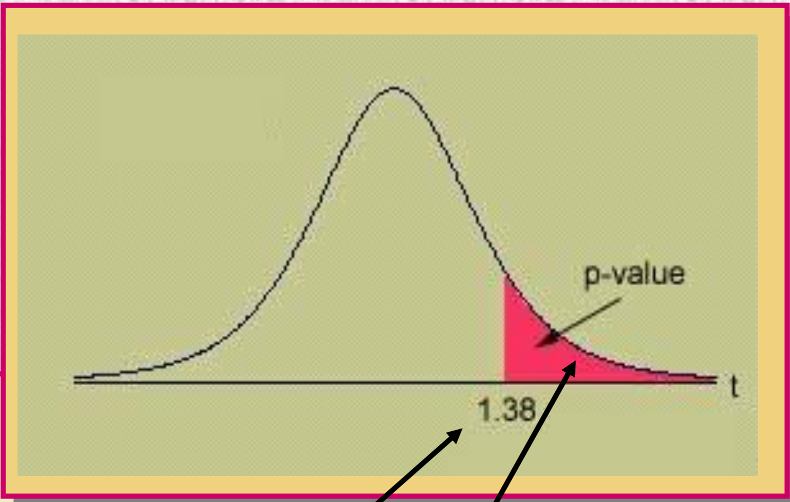
Since the observed value of $t = 1.38$ is smaller than $t_{.10} = 1.476$,

p-value $> .10$.

The exact *p*-value

- You can get the exact *p*-value using some calculators or a computer.

p-value = 0.113 which is greater than 0.10 as we approximated using *t* –table.



One-Sample T: Times

```
Test of mu = 15 vs mu > 15

Variable          N      Mean     StDev   SE Mean
Times             6     19.17      7.39      3.02

Variable       95.0% Lower Bound        T         P
Times                      13.09     1.38     0.113
```

Testing the Difference between Two Means (Case 1)

As in Chapter 9, independent random samples of size n_1 and n_2 are drawn from populations 1 and 2 with means μ_1 and μ_2 and variances σ_1^2 and σ_2^2.

- To test:

H_0: $\mu_1 - \mu_2 = D_0$ versus H_a: one of three

where D_0 is some hypothesized difference, usually 0.

Testing the Difference between Two Means (Case 1)

- The test statistic used in Chapter 9

$$z \approx \frac{\bar{x}_1 - \bar{x}_2}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}}$$

  does not have either a *z* or a *t* distribution, and cannot be used for small-sample inference.
- We need to make one more assumption, that **the population variances, although unknown, are equal.**

Testing the Difference between Two Means (Case 1)

The two assumptions:

1. The two samples are randomly and independently selected from normal populations.

2. Population variances are equal. That is $\sigma_1^2 = \sigma_2^2$.

Testing the Difference between Two Means (Case 1)

Instead of estimating each population variance separately, we estimate the common variance with

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

s^2 *is an unbiased estimator of both* $\sigma_1^2 = \sigma_2^2$ *and called a pooled sample variance.*

And the resulting test statistic,

$$t = \frac{\bar{x}_1 - \bar{x}_2 - D_0}{\sqrt{s^2\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}}$$

has a t distribution with n_1+n_2-2 degrees of freedom.

Pooled t-test

Estimating the Difference between Two Means (Case 1)

A 100(1-α)% confidence interval for μ_1-μ_2 is.

$$(\bar{x}_1 - \bar{x}_2) \pm t_{\alpha/2} \sqrt{s^2\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}$$

with $s^2 = \dfrac{(n_1 - 1)s_1^2 + (n_2 - 1)s_2^2}{n_1 + n_2 - 2}$

and the critical value $t_{\alpha/2}$ is based on $(n_1 + n_2 - 2)$ degrees of freedom.

Example 2

Two training procedures are compared by measuring the time that it takes trainees to assemble a device. A different group of trainees are taught using each method. Is there a difference between the mean times in the two methods? Test using $\alpha = 0.01$. Assume that the times are normally distributed.

Time to Assemble	**Method 1**	**Method 2**
Sample size	10	12
Sample mean	35	31
Sample Std. Dev.	4.9	4.5

$$H_0 : \mu_1 - \mu_2 = 0$$

$$H_a : \mu_1 - \mu_2 \neq 0$$

Test Statistic :

$$t = \frac{\bar{x}_1 - \bar{x}_2 - 0}{\sqrt{s^2\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}}$$

Example 2

- Solve this problem by approximating the *p*-value using Table 4.

Time to Assemble	Method 1	Method 2
Sample size	10	12
Sample mean	35	31
Sample Std. Dev.	4.9	4.5

Calculate:

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

$$= \frac{9(4.9^2) + 11(4.5^2)}{20} = 21.942$$

Test statistic:

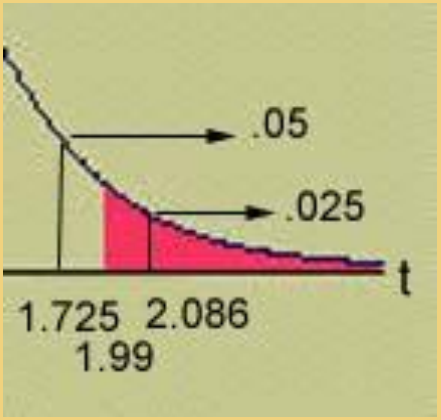
$$t = \frac{35 - 31}{\sqrt{21.942\left(\frac{1}{10} + \frac{1}{12}\right)}}$$

$$= 1.99$$

Example 2

$$p-\text{value} = P(t < -1.99) + P(t > 1.99)$$

$$\begin{aligned} df &= n_1 + n_2 - 2 \\ &= 10 + 12 - 2 \\ &= 20 \end{aligned}$$

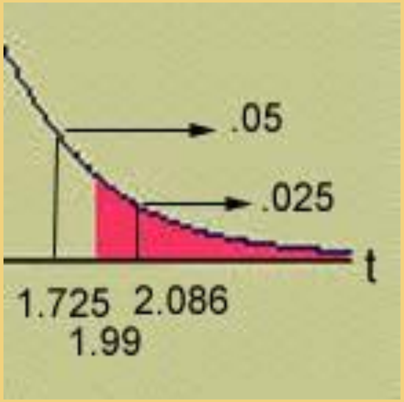


$P(t > 1.99)$ is between 0.025 and 0.05.

df	$t_{.100}$	$t_{.050}$	$t_{.025}$	$t_{.010}$	$t_{.005}$	df
19	1.328	1.729	2.093	2.539	2.861	19
20	1.325	1.725	2.086	2.528	2.845	20

Example 2

$$p - \text{value} = P(t < -1.99) + P(t > 1.99)$$



Thus,

$$\mathbf{2(0.025) < \textit{p}\text{-value} < 2(0.05)}$$

$$\mathbf{0.05 < \textit{p}\text{-value} < 0.10}$$

Since the *p*-value is greater than $\alpha = .01$, H_0 is not rejected.

There is insufficient evidence to indicate that Methods 1 and 2 yield different mean times.

Example 2

A 99% confidence interval for $\mu_1-\mu_2$ is.

$$(\bar{x}_1-\bar{x}_2)\pm t_{\alpha/2}\sqrt{s^2\left(\frac{1}{n_1}+\frac{1}{n_2}\right)}$$

$$= 4\pm 2.845\sqrt{21.942\left(\tfrac{1}{10}+\tfrac{1}{12}\right)}$$

$$= 4\pm 5.706$$

Based on a 99% CI, $\mu_1-\mu_2$ is estimated to be between -1.706 and 9.706 minutes.

Because 0 is contained in this interval, we conclude that the mean times to assemble the device using Methods 1 and 2 are not different.

Testing the Difference between Two Means (Case 1)

How can you tell if the equal variance assumption is reasonable?

Rule of Thumb :

If the ratio, $\dfrac{\text{larger } s^2}{\text{smaller } s^2} \le 3$,

the equal variance assumption is reasonable.

If the ratio, $\dfrac{\text{larger } s^2}{\text{smaller } s^2} > 3$,

use an alternative test statistic.

Testing the Difference between Two Means (Case 1)

If the population variances cannot be assumed equal, the test statistic

$$t \approx \frac{\bar{x}_1 - \bar{x}_2}{\sqrt{\dfrac{s_1^2}{n_1} + \dfrac{s_2^2}{n_2}}}$$

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

has an approximate *t* distribution with degrees of freedom given above. This is most easily done by computer.

Testing the Difference between Two Means (Case 2)
The Paired-Difference Test

- Sometimes the assumption of independent samples is intentionally violated, resulting in a matched-pairs or paired-difference test.
- By designing the experiment in this way, we can eliminate unwanted variability in the experiment by analyzing only the differences,

$$d_i = x_{1i} - x_{2i}$$

to see if there is a difference in the two population means, μ_1 and μ_2.

Example 3

Salesperson	1	2	3	4	5	6
Before (B)	12	18	25	9	14	16
After (A)	18	24	24	14	19	20

A company wanted to know if attending a course on "how to be a successful salesperson" can increase the average sales of its employees (in 10,000 baht). The company sent six of its salespersons to attend this course. The table gives the one-week sales of these salespersons before and after they attended this course.

At 1% significance level, can you conclude that the mean weekly sales increase as a result of attending this course? Assume that the population of paired difference has a normal distribution.

Example 3

Salesperson	1	2	3	4	5	6
Before (B)	12	18	25	9	14	16
After (A)	18	24	24	14	19	20

$$\mathrm{H}_0 : \mu_B - \mu_A = 0$$
$$\mathrm{H}_a : \mu_B - \mu_A < 0$$

Equivalently,

$$\mathrm{H}_0 : \mu_d = 0$$
$$\mathrm{H}_a : \mu_d < 0$$

where $\mu_d = \mu_B - \mu_A$

The samples are not independent. The pairs of responses are linked because measurements are taken on the same salesperson.

The Paired-Difference Test

To test $H_0: \mu_1 - \mu_2 = 0$ we test $H_0: \mu_d = 0$

using the test statistic

$$t = \frac{\bar{d} - 0}{s_d / \sqrt{n}}$$

where n = number of pairs,

$\bar{d}$ and s_d are the mean and standard deviation of the differences, d_i.

Use the p-value or a rejection region based on a t-distribution with $df = n - 1$.

The Paired-Difference Test

A 100(1-α)% confidence interval for μ_d is

$$\bar{d} \pm t_{\alpha/2} \frac{s_d}{\sqrt{n}}$$

The critical value $t_{\alpha/2}$ is based on n-1 degrees of freedom.

Example 3

Salesperson	1	2	3	4	5	6
Before (B)	12	18	25	9	14	16
After (A)	18	24	24	14	19	20
Difference, d_i	-6	-6	1	-5	-5	-4

$$H_0 : \mu_d = 0$$
$$H_a : \mu_d < 0$$

$$\text{Calculate } \bar{d} = \frac{\sum d_i}{n} = \frac{-25}{6} = -4.17$$

$$s_d = \sqrt{\frac{\sum d_i^2 - \frac{\left(\sum d_i\right)^2}{n}}{n-1}} = \sqrt{\frac{139 - \frac{(-25)^2}{6}}{5-1}} = 2.63944$$

$$\text{Test Statistic}: \quad t = \frac{\bar{d} - 0}{s_d / \sqrt{n}} = \frac{-4.17}{2.63944 / \sqrt{6}}$$
$$= -3.870$$

Example 3

Salesperson	1	2	3	4	5	6
Before (B)	12	18	25	9	14	16
After (A)	18	24	24	14	19	20
Difference, d_i	-6	-6	1	-5	-5	-4

Rejection region: At $\alpha = 0.01$, reject H_0 if

$t < -t_{0.01}$ with $n-1 = 5$ *df*.

$t < -3.365$

Conclusion: Since $t = -3.87$, H_0 is rejected.

Consequently, we conclude that the mean weekly sales increase as a result of this course.

Example 3

Salesperson	1	2	3	4	5	6
Before (B)	12	18	25	9	14	16
After (A)	18	24	24	14	19	20
Difference, d_i	-6	-6	1	-5	-5	-4

Find a 95% confidence interval for the difference between the mean weekly sales before versus after.

Some Notes

- You can construct a 100(1-α)% confidence interval for a paired experiment using

$$\boxed{\bar{d} \pm t_{\alpha/2} \frac{s_d}{\sqrt{n}}}$$

- Once you have designed the experiment by pairing, you MUST analyze it as a paired experiment. If the experiment is not designed as a paired experiment in advance, do not use this procedure.