Introduction to Probability and Statistics Eleventh Edition

Chapter 10
Inference from Small Samples

Introduction

- Our hypothesis tests so far have assumed two things:

1. The sampling distribution is approximately normal (for large n we often use the CLT).

2. We know the population standard deviation σ.

- If either of these is not true then, strictly speaking, the methods shown so far do not apply.

The Sampling Distribution of the Sample Mean (non-normal population)

- If n is large, we assume the sampling distribution of $\bar{x}$ is approximately normal (by the CLT). So the test statistic

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

will approximately follow a standard normal distribution.

- But if n is not large, then this test statistic **may not be approximately normal**, so we can't use the table to calculate probabilities like we want to.

The Sampling Distribution of the Sample Mean (normal population)

- If the original population is normal, then the CLT says the sampling distribution of sample means will also be normal.
- So we can do a z-test using

$$z = \frac{\bar{x} - \mu}{\frac{\sigma}{\sqrt{n}}}$$

- However, we still need to know the population standard deviation σ otherwise we can't calculate the test statistic z.
- We've been cheating by using the sample standard deviation s in place of σ.

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

The Sampling Distribution of the Sample Mean (normal population)

- This cheat kind of works, but to get better results we need to make some changes.
- The problem is that the standard deviation of a sample (s) is not a perfect estimator for the standard deviation of the population we take it from (σ).
- This is mainly a problem for small sample sizes.
- This means the test statistic

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

often does not quite follow a normal distribution.

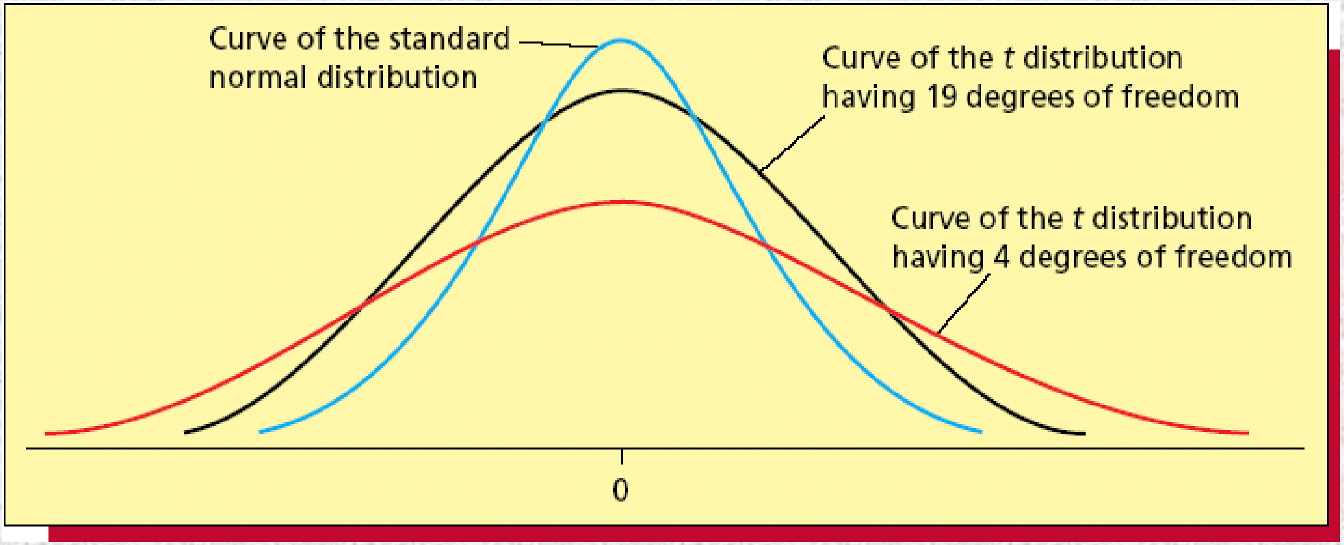
Student's *t* Distribution

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

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

- 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, *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 $\boldsymbol{a}$ to find t_a, the value of t with area $\boldsymbol{a}$ to its right.

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. The only difference is we use t not z.
- 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

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.

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

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

Example 1

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

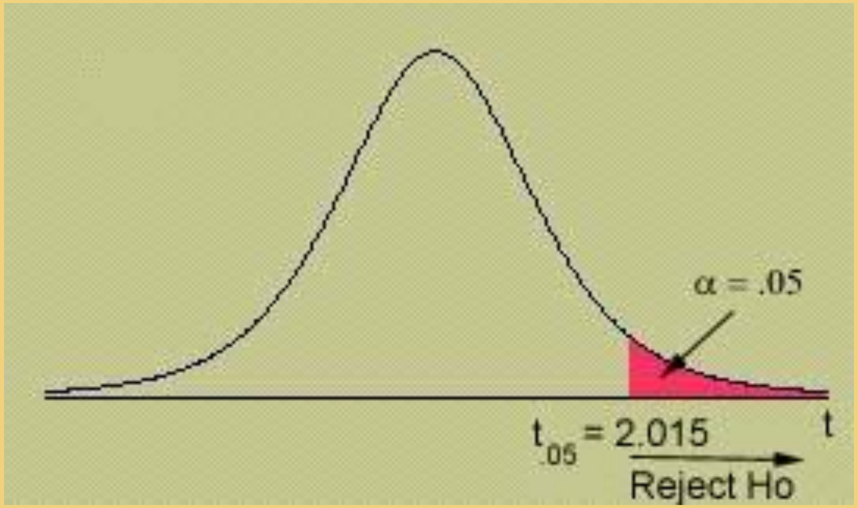
Calculate the test statistic and find the rejection region for α =.05.

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.

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

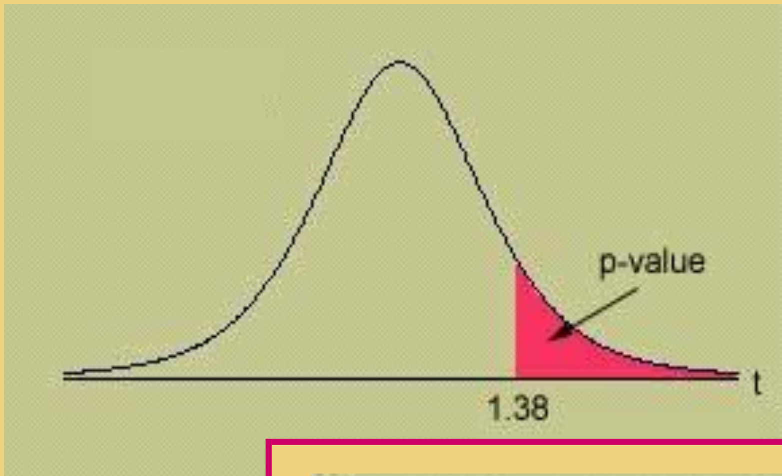
Test statistic : $t = 1.38$

Rejection Region: Reject H_0 if $t > 2.015$

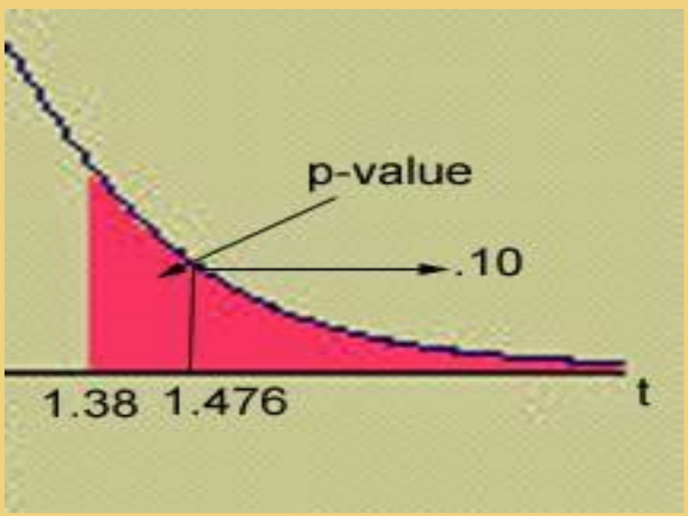
Conclusion: For our example, $t = 1.38$ does not fall in the rejection region and H_0 is not rejected. There is insufficient evidence to indicate that the average activation time is greater than 15.

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



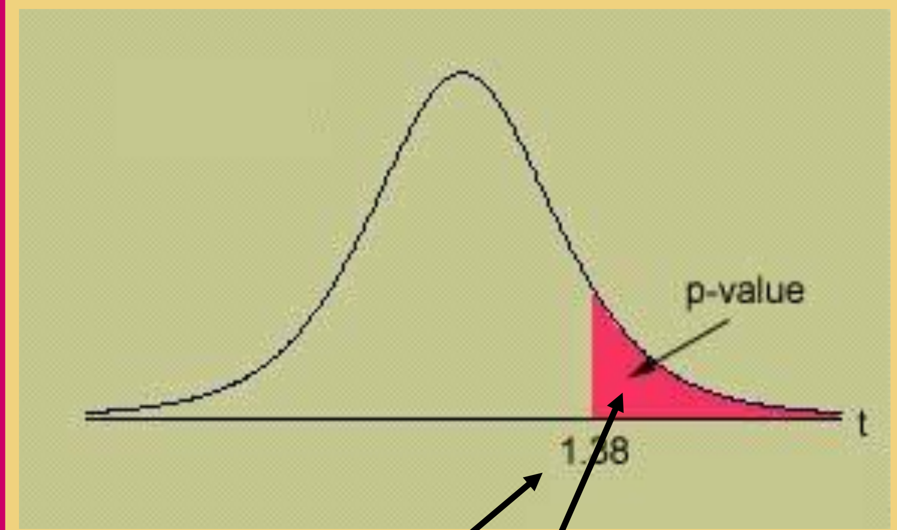
Since the observed value of $t = 1.38$ is smaller than $t_{.10} = 1.476$,

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
```

Class Activity 13

1. The telephone company is interested in measuring the average daily usage in minutes for household telephones in a specific area. Suppose that a random sample of nine households were sampled on random days, producing the following times (in minutes): 35, 59, 42, 44, 31, 46, 24, 56, 50

   a. Estimate the daily usage using a 90% confidence interval. You can assume usage follows a normal distribution.

   b. Suppose that the statewide average for households has been found to be 45 minutes. Does this data provide evidence to indicate that the average in this area differs from the statewide average? Use $\alpha=0.05$.

Testing the Difference between Two Means

Independent random samples of size n_1 and n_2 are drawn from normal populations A and B with means μ_1 and μ_2 and standard deviations σ_1 and σ_2.

- We want to test whether the difference in population means $\mu_1 - \mu_2$ is different in some way from some hypothesized value D_0.
- We could need either left, right, or two tailed test, depending on our hypothesis.

Testing the Difference between Two Means

- We usually don't know σ_1 or σ_2.
- Previously we cheated and used the sample standard deviations

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

- This isn't always very accurate, and if we want to be good statisticians we should be careful.

Testing the Difference between Two Means

Suppose we know the variance of the two populations is very similar.

We combine the samples to estimate the variance of the populations using the formula

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

Pooled sample variance

We use this estimate to calculate a test statistic using

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

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

Pooled t-test

Estimating the Difference between Two Means

You can also create a $100(1-\alpha)\%$ confidence interval for $\mu_1-\mu_2$.

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

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

Remember the three assumptions:

1. **Original populations are <u>normal</u>.**
2. **Samples are <u>random</u> and <u>independent</u>.**
3. **Population <u>variances are approximately equal</u>.**

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 in the two methods? Use $\alpha = .01$.

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 :

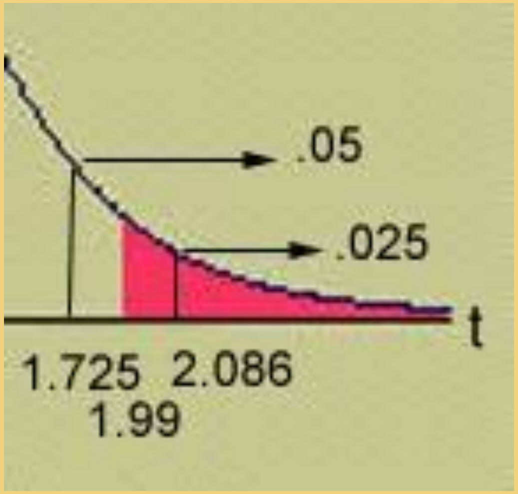
$$t = \frac{35 - 31}{\sqrt{21.942\left(\frac{1}{10} + \frac{1}{12}\right)}}$$

$$= 1.99$$

Example 2

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

$df = n_1 + n_2 - 2 = 10 + 12 - 2 = 20$



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

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

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

Since the *p*-value is greater than $\alpha = .01$, H_0 is not rejected. There is insufficient evidence to indicate a difference in the population means.

Testing the Difference between Two Means

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

Rule of Thumb :

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

the equal variance assumption is reasonable.

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

use an alternative test statistic.

Class Activity 13

2. In the process of deciding whether to close a civic health center, a random sample of 25 people who had visited the center at least once was chosen and each person asked whether the center should be closed. Also the distance between each person's home and the health center was recorded. 16 people were in favor of not closing. For these people, the average distance from the center was 5.2 miles with a variance of 2.8 miles. The 9 people in favor of closing lived at an average of 8.7 miles from the center with a variance of 5.3 miles. Does the data indicate that there is a significant difference in mean distance to the health center for these two groups? Use $\alpha = 0.05$.

Class Activity 13

3. Estimate the difference in mean distance to the health center for the two groups in Exercise 2 with a 95% confidence interval.

Testing the Difference between Two Means

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.

The Paired-Difference Test

- Sometimes the assumption of <u>independent</u> samples is intentionally <u>violated</u>, 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, $\mu_1 - \mu_2$.

Example 3

Car	1	2	3	4	5
Type A	10.6	9.8	12.3	9.7	8.8
Type B	10.2	9.4	11.8	9.1	8.3

- One Type A and one Type B tire are randomly assigned to each of the rear wheels of five cars. Compare the average tire wear for types A and B using a test of hypothesis. Assume wear on tires follows a normal distribution.

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

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

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 $df = n - 1$.

Example

Car	1	2	3	4	5
Type A	10.6	9.8	12.3	9.7	8.8
Type B	10.2	9.4	11.8	9.1	8.3
Difference	**0.4**	**0.4**	**0.5**	**0.6**	**0.5**

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

$$\text{Calculate } \bar{d} = \frac{\sum d_i}{n} = .48$$

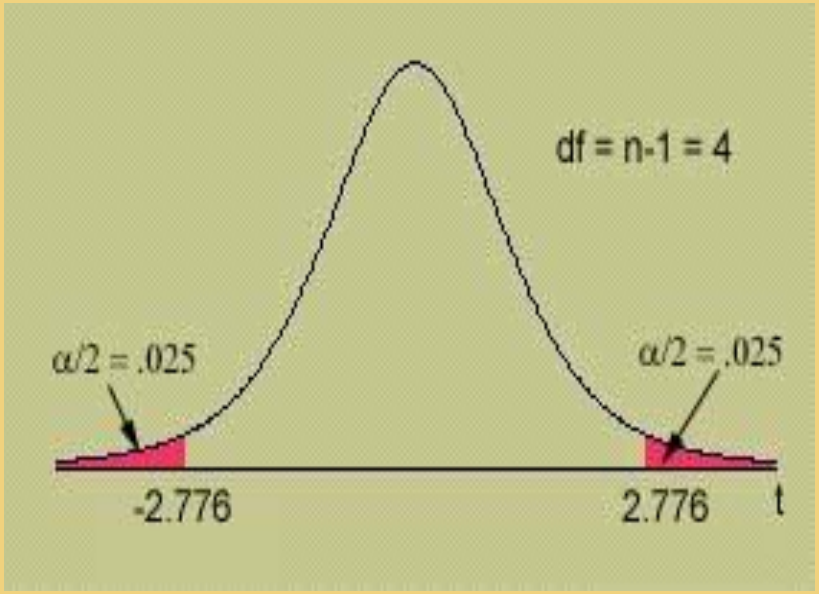
$$s_d = \sqrt{\frac{\sum d_i^2 - \frac{\left(\sum d_i\right)^2}{n}}{n-1}} = .0837$$

Test statistic:

$$t = \frac{\bar{d} - 0}{s_d / \sqrt{n}} = \frac{.48 - 0}{0.0837 / \sqrt{5}} = 12.8$$

Example

Car	1	2	3	4	5
Type A	10.6	9.8	12.3	9.7	8.8
Type B	10.2	9.4	11.8	9.1	8.3
Difference	**0.4**	**0.4**	**0.5**	**0.6**	**0.5**



Rejection region: Reject H_0 if

$t > 2.776$ or $t < -2.776$.

Conclusion: Since $t = 12.8$, H_0 is rejected. There is a difference in the average tire wear for the two types of tires.

Some Notes

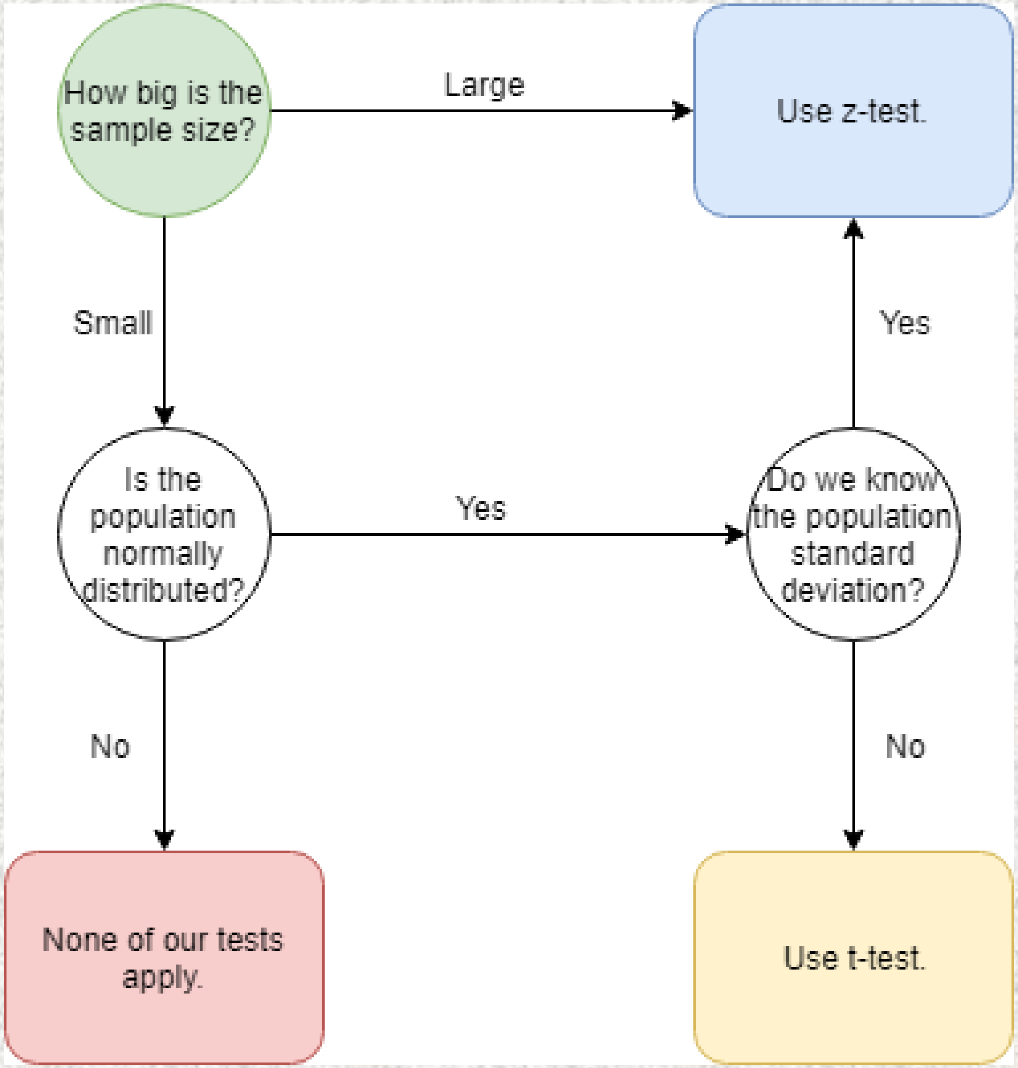
You can construct a 100(1-α)% confidence interval for a paired experiment using

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

Use a paired difference test if and only if the points in the two samples are naturally paired.

Paired-difference testing can be done with either a t-statistic or a z-statistic, depending on the sample size.

z vs t: which test to use?



Class Activity 13

4. Students are taught the same syllabus using two different methods. An experimenter randomly chose two groups of 18 students each, and assigned one group to use method I and the second to use method II. After the courses a quiz was given to each group and their grades recorded. Do the following data indicate a difference in the mean quiz scores for the two methods? Find the approximate *p-value* and interpret your results.

	$\bar{x}$	s^2
Presentation I	81.7	23.2
Presentation II	77.2	19.8

Class Activity 13

5. The owner of a small manufacturing plant is considering a change in salary base by replacing an hourly wage structure with a per-unit rate. Before arriving at any decision, she forms 10 pairs of workers so that within each pair the two workers have produced about the same number of items per day and their work has been of comparable quality. From each pair, one worker is randomly selected to be paid as usual and the other is to be paid on a per-unit basis. In addition to the number of items produced, a cumulative quality score for the items produced is kept for each worker. The quality scores follow. (A high score is indicative of high quality.)

Class Activity 13

5. cont.

	Rate	
Pair	*Per Unit*	*Hourly*
1	86	91
2	75	77
3	87	83
4	81	84
5	65	68
6	77	76
7	88	89
8	91	91
9	68	73
10	79	78

Do these data indicate that the average quality for the per-unit production is significantly lower than that based on an hourly wage? What *p-value* would you report?

Class Activity 13

6. Refer to Exercise 5. The following data represent the average number of items produced per worker, based on one week's production records:

	Rate	
Pair	*Per Unit*	*Hourly*
1	35.8	31.2
2	29.4	27.6
3	31.2	32.2
4	28.6	26.4
5	30.0	29.0
6	32.6	31.4
7	36.8	34.2
8	34.4	31.6
9	29.6	27.6
10	32.8	29.8

a. Estimate the mean difference in average daily output for the two pay scales with a 95% confidence interval.

b. Test the hypothesis that a per-unit pay scale increases production at the .05 level of significance.

Videos

- Introduction to the t Distribution (non-technical) **https://www.youtube.com/watch?v=Uv6nGIgZMVw**
- Hypothesis tests on one mean: t test or z test? **https://www.youtube.com/watch?v=vw2IPZ2aD-c**
- Pooled-Variance t Tests and Confidence Intervals: An Example **https://www.youtube.com/watch?v=Q526z1mz4Sc**
- An Introduction to Paired-Difference Procedures **https://www.youtube.com/watch?v=tZZt8f8URKg**