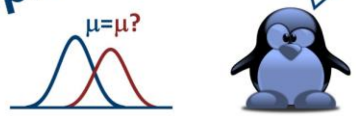
The two sample t-test

To see if means are different

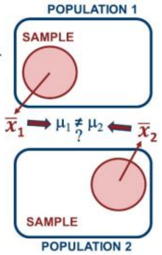


Comparing population means

We want to know if the population means differ.
We can't measure the populations.
We take random samples.
We calculate sample means.

The sample means are estimates of the population means, but sampling error makes them inexact.

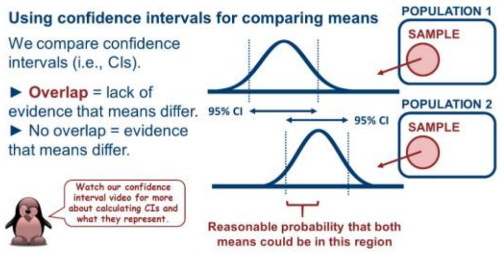
What are the chances the population means are the same (i.e., H_0), based on how much the sample means differ from one another?



Using confidence intervals for comparing means

We compare confidence intervals (i.e., CIs).

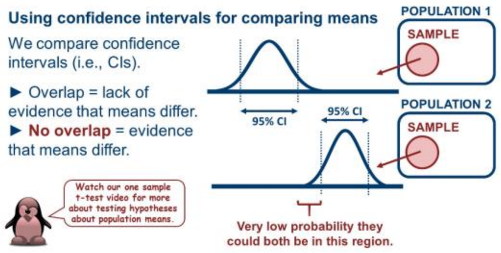
▶ **Overlap** = lack of evidence that means differ.
▶ No overlap = evidence that means differ.



Using confidence intervals for comparing means

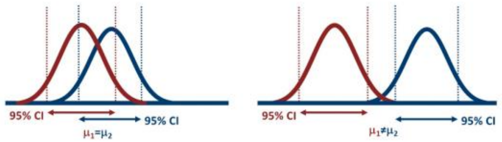
We compare confidence intervals (i.e., CIs).

▶ Overlap = lack of evidence that means differ.
▶ **No overlap** = evidence that means differ.



The t test is a comparison of confidence intervals

We could calculate both confidence intervals and see if they overlap. Instead we usually calculate a single t or Z value for the difference and compare it to zero (i.e., when H_0 is true).



There is also a paired t-test

If every value in each data set has a specific partner, then the set of individual differences should have a mean of zero. One set of values with an μ_0=0 can be tested with a one-sample t-test.

Set 1	Set 2	Diff.
$x1_1$	$x2_1$	D_1
$x1_2$	$x2_2$	D_2
$x1_3$	$x2_3$	D_3
...	...	
$x1_n$	$x2_n$	D_n

$$t_{calc} = \frac{\bar{D} - 0}{\frac{s_D}{\sqrt{n}}}$$

Requires paired values.
e.g., twin studies or before/after

What is the combined standard error?

$$t_{calc} = \frac{\bar{x}_1 - \bar{x}_2}{combined\ SE}$$

Recall: $SE = \frac{s}{\sqrt{n}}$

If population variances are known

$$SE = \sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}$$

If population variances are unknown, but equal

$$SE = \sqrt{s_p^2\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}$$

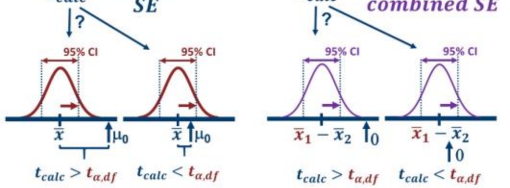
$$s_p^2 = \frac{(n_1 - 1)s_1^2 + (n_2 - 1)s_2^2}{(n_1 - 1) + (n_2 - 1)}$$

If population variances are unknown, and not equal

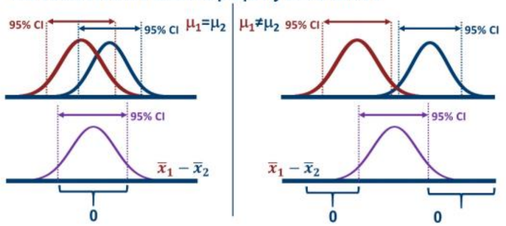
$$SE = \sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}$$

Analogy between one-sample and two-sample t-test

$$t_{calc} = \frac{\bar{x} - \mu_0}{SE} \qquad t_{calc} = \frac{\bar{x}_1 - \bar{x}_2}{combined\ SE}$$



The t test uses the additive property of variances



The t test uses the additive property of variances

Theorem: For two uncorrelated or independent data sets A and B, the variance of the combined data set is the sum of the separate variances:

$$\sigma^2_{A\pm B} = \sigma^2_A + \sigma^2_B$$

Implication: If the means of two populations are equal, then the difference between two samples should be zero and the combined variance will be sum of the original two.

Application: A t-test where H_0 is that μ_1-μ_2=0 can be used to see if the means of two populations differ.

Degrees of freedom for the two-sample t-tests

Unpaired two sample Z-test

$$Z_{calc} = \frac{\bar{x}_1 - \bar{x}_2}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}}$$

$$df = n_1 + n_2 - 2$$

Unpaired two sample homoscedastic "<u>student's</u>" t-test

$$t_{calc} = \frac{\bar{x}_1 - \bar{x}_2}{\sqrt{s_p^2\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}}$$

$$df = n_1 + n_2 - 2$$

Degrees of freedom for the two-sample t-tests

Paired two sample t-test

$$t_{calc} = \frac{\bar{D} - 0}{\frac{s_D}{\sqrt{n}}}$$

$$df = n - 1$$

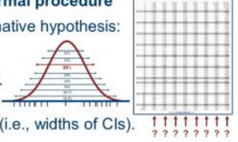
Unpaired two sample heteroscedastic "<u>student's</u>" t-test

$$t_{calc} = \frac{\bar{x}_1 - \bar{x}_2}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}} \qquad df = \frac{\left(\left(\frac{s_1^2}{n_1}\right) + \left(\frac{s_2^2}{n_2}\right)\right)^2}{\frac{\left(\frac{s_1^2}{n_1}\right)^2}{n_1 - 1} + \frac{\left(\frac{s_2^2}{n_2}\right)^2}{n_2 - 1}}$$

These are the two most important t-tests

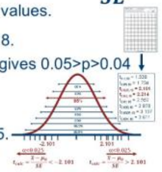
The two-tailed two-sample t test formal procedure

▶ Create a null hypothesis and alternative hypothesis:
H_0: $\mu_1 = \mu_2$
H_A: $\mu_1 \neq \mu_2$
▶ Calculate t_{calc} $\quad t_{calc} = \frac{\bar{x}_1 - \bar{x}_2}{SE}$
▶ Compare t_{calc} to various t_{crit} values (i.e., widths of CIs).
▶ Determine probability, **<u>p value</u>**, of seeing t_{calc} as extreme as we do.
▶ Decide to "reject H_0" or "fail to reject H_0" based on the p value.
H_0: $\mu_1 = \mu_2$ consistent with non-small p values.
H_A: $\mu_1 \neq \mu_2$ would give us small p values.



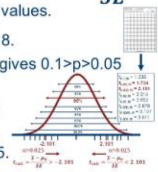
The t test practical procedure

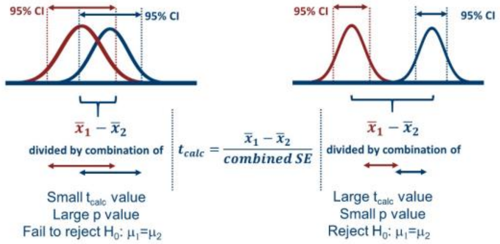
▶ Create a null hypothesis and alternative hypothesis:
H_0: $\mu_1 = \mu_2$ and H_A: $\mu_1 \neq \mu_2$ $\quad t_{calc} = \frac{\bar{x}_1 - \bar{x}_2}{SE}$
▶ Calculate t_{calc} and compare t_{calc} to various t_{crit} values.
▶ Determine the p value - e.g., t_{calc}=**2.2** for df=18.
Use table: $t_{0.025,18}$=2.101<**2.2**<2.214=$t_{0.02,18}$ gives 0.05>p>0.04
Use computer: calculation gives p=0.041
▶ Use the small p value to "reject H_0"
H_0: $\mu_1 = \mu_2$ not consistent with p=0.041<0.05.
H_A: $\mu_1 \neq \mu_2$ is consistent with p=0.041<0.05.



The t test practical procedure

▶ Create a null hypothesis and alternative hypothesis:
H_0: $\mu_1 = \mu_2$ and H_A: $\mu_1 \neq \mu_2$ $\quad t_{calc} = \frac{\bar{x}_1 - \bar{x}_2}{SE}$
▶ Calculate t_{calc} and compare t_{calc} to various t_{crit} values.
▶ Determine the p value - e.g., t_{calc}=**2.0** for df=18.
Use table: $t_{0.05,18}$=1.734<**2.0**<2.101=$t_{0.025,18}$ gives 0.1>p>0.05
Use computer: calculation gives p=0.061
▶ Can't use the moderate p value to "reject H_0"
H_0: $\mu_1 = \mu_2$ is consistent with p=0.061>0.05.
H_A: $\mu_1 \neq \mu_2$ not consistent with p=0.061>0.05.





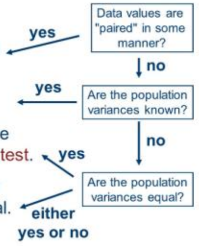
Which t-test to use?

If the individual data values are paired in some way. Use paired t-test.

If we know σ. Use Z-test (unrealistic).

If we don't know σ, but we know they're equal. Use unpaired homoscedastic t-test.

If we don't know σ, and we don't know (or don't want to assume) they're equal. Use unpaired heteroscedastic t-test.

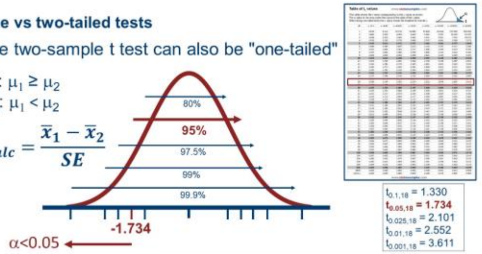


One vs two-tailed tests

The two-sample t test can also be "one-tailed"

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

$$t_{calc} = \frac{\bar{x}_1 - \bar{x}_2}{SE}$$

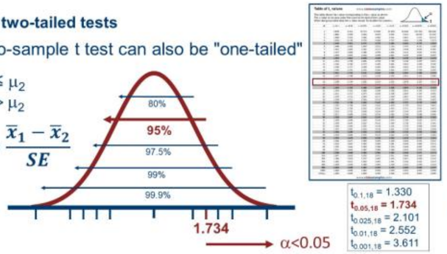


One vs two-tailed tests

The two-sample t test can also be "one-tailed"

H_0: $\mu_1 \leq \mu_2$
H_A: $\mu_1 > \mu_2$

$$t_{calc} = \frac{\bar{x}_1 - \bar{x}_2}{SE}$$



Statistically significant

The use of p=0.05 (i.e., 5%) as a threshold for deciding to reject null hypothesis is arbitrary, but is the standard.

<u>Statistically significant</u>: A test has returned a p value less than the threshold and the null hypothesis has been rejected.

▶ If sample means of 18 and 20 are *significantly different*, (i.e., p<0.05), then we reject the null hypothesis that the population means are equal (we can also describe direction).
▶ If sample means of 18 and 20 are *not significantly different*, (i.e., p>0.05), then we fail to reject the null hypothesis that the population means are equal (we may assume they're equal).

StatsExamples.com