

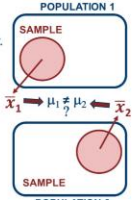
Two sample unpaired t-test: examples

Review: two-sample "student's" t-test

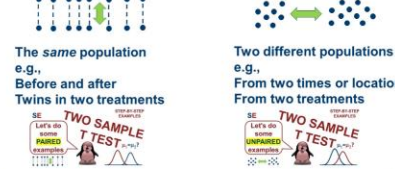
We want to know if the population means differ. We can't measure the populations so we take random samples and calculate sample means.

The sample means estimate the population means, but sampling error makes them imprecise.

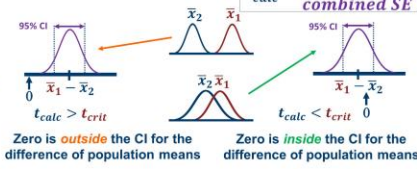
What are the chances the population means are the same (i.e., H_0), based on how much the sample means differ from one another?



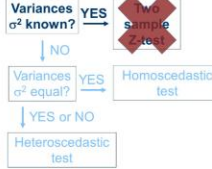
Two-sample "student's" t-test: paired vs unpaired data values



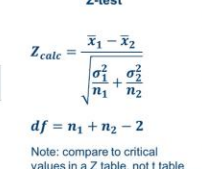
Two-sample t-test equation concept



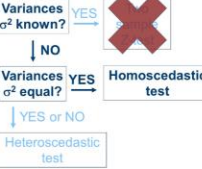
Two-sample t-test choice



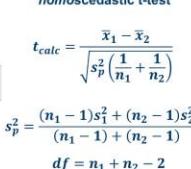
Unpaired two sample Z-test



Two-sample t-test choice



Unpaired two sample homoscedastic t-test



Example #1 - homoscedastic test

Reef A: $\bar{x}_A = 31.00, s_A = 3.200$
Reef B: $\bar{x}_B = 29.00, s_B = 4.444$
 $df = 19$ $t_{calc} = 2.3514$

$|t_{calc}| > |t_{crit}|?$
 $2.3514 > 2.093?$

The mean masses are significantly different
Q1: Which is population has the larger mean?
Q2: What is the p value?

Example #1 - homoscedastic test

Reef A: $\bar{x}_A = 31.00, s_A = 3.200$ $df = n_A + n_B - 2 = 11 + 10 - 2 = 19$
Reef B: $\bar{x}_B = 29.00, s_B = 4.444$

$s_p^2 = \frac{(n_A - 1)s_A^2 + (n_B - 1)s_B^2}{(n_A - 1) + (n_B - 1)} = \frac{(11 - 1)3.200^2 + (10 - 1)4.444^2}{(11 - 1) + (10 - 1)} = 3.78947$

$t_{calc} = \frac{\bar{x}_A - \bar{x}_B}{\sqrt{s_p^2 \left(\frac{1}{n_A} + \frac{1}{n_B} \right)}} = \frac{31.00 - 29.00}{\sqrt{3.78947 \left(\frac{1}{11} + \frac{1}{10} \right)}} = \frac{2.00}{0.85056} = 2.3514$

Example #1 - masses of fish

Masses of fish caught at each of two reefs:
Reef A: 32, 32, 31, 32, 33, 27, 29, 30, 31, 33, 31
Reef B: 30, 27, 28, 27, 28, 32, 29, 27, 33, 29

Are the mean masses of the fish populations the same?
With what degree of confidence do we decide?
First step, calculate sample means and variances:

Reef A: $\bar{x}_A = 31.00, s_A^2 = 3.200$
Reef B: $\bar{x}_B = 29.00, s_B^2 = 4.444$

The 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} .
- Compare t_{calc} to various t_{crit} values (i.e., widths of CIs).
- Determine probability, **p value**, of seeing t_{calc} as extreme as we do.
- Decide to "fail to reject H_0 " or "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.

Example #1 - homoscedastic test

Reef A: $\bar{x}_A = 31.00, s_A = 3.200$
Reef B: $\bar{x}_B = 29.00, s_B = 4.444$
 $df = 19$ $t_{calc} = 2.3514$

$|t_{calc}| > |t_{crit}|?$
 $2.3514 > 2.093?$

"The mean mass of fish at reef A is significantly larger than the mean mass of fish at reef B ($0.02 < p < 0.04$).

Example #1 - heteroscedastic test

Reef A: $\bar{x}_A = 31.00, s_A = 3.200$
Reef B: $\bar{x}_B = 29.00, s_B = 4.444$

$t_{calc} = \frac{\bar{x}_A - \bar{x}_B}{\sqrt{\frac{s_A^2}{n_A} + \frac{s_B^2}{n_B}}} = \frac{31.00 - 29.00}{\sqrt{\frac{3.200^2}{11} + \frac{4.444^2}{10}}} = \frac{2.00}{0.85056} = 2.3323$

$df = \frac{\left(\frac{s_A^2}{n_A} + \frac{s_B^2}{n_B} \right)^2}{\frac{s_A^2}{n_A} + \frac{s_B^2}{n_B}} = \frac{\left(\frac{3.200^2}{11} + \frac{4.444^2}{10} \right)^2}{\frac{3.200^2}{11} + \frac{4.444^2}{10}} = \frac{0.54074}{0.03041} = 17.814 = 17$

$|t_{calc}| > |t_{crit}|?$
 $2.3323 > 2.110?$

The mean masses are significantly different
Q1: What is the p value?

Example #1 - heteroscedastic test

Reef A: $\bar{x}_A = 31.00, s_A = 3.200$
Reef B: $\bar{x}_B = 29.00, s_B = 4.444$
 $df = 17$ $t_{calc} = 2.3323$

$|t_{calc}| > |t_{crit}|?$
 $2.3323 > 2.110?$

The mean masses are significantly different
Q1: What is the p value?

Example #1 - heteroscedastic test

Reef A: $\bar{x}_A = 31.00, s_A = 3.200$
Reef B: $\bar{x}_B = 29.00, s_B = 4.444$
 $df = 17$ $t_{calc} = 2.3323$

$|t_{calc}| > |t_{crit}|?$
 $2.3323 > 2.110?$

"The mean mass of fish at reef A is significantly larger than the mean mass of fish at reef B ($0.02 < p < 0.04$).

Example #3 - Cracks at construction sites

Consider number of cracks each of two locations:
Site 1: 13, 11, 18, 6, 13, 5
Site 2: 10, 4, 6, 1, 6, 12, 7, 2

Are the mean numbers of cracks the same?
With what degree of confidence do we decide?
First step, calculate sample means and variances:

Site 1: $\bar{x}_1 = 11.00, s_1^2 = 23.600$
Site 2: $\bar{x}_2 = 6.00, s_2^2 = 14.000$

Example #2 - heteroscedastic test

Old: $\bar{x}_{old} = 16.00, s_{old}^2 = 5.1429$
New: $\bar{x}_{new} = 14.00, s_{new}^2 = 6.8571$
 $df = 13$ $t_{calc} = 1.6330$

$|t_{calc}| > |t_{crit}|?$
 $1.6330 > 2.160?$

"The mean recovery times for the old and new procedure are not significantly different ($0.1 < p < 0.2$).

Example #2 - heteroscedastic test

Old: $\bar{x}_{old} = 16.00, s_{old}^2 = 5.1429$
New: $\bar{x}_{new} = 14.00, s_{new}^2 = 6.8571$

$t_{calc} = \frac{\bar{x}_{old} - \bar{x}_{new}}{\sqrt{\frac{s_{old}^2}{n_{old}} + \frac{s_{new}^2}{n_{new}}}} = \frac{16.00 - 14.00}{\sqrt{\frac{5.1429}{11} + \frac{6.8571}{10}}} = \frac{2.00}{0.85056} = 1.6330$

$df = \frac{\left(\frac{s_{old}^2}{n_{old}} + \frac{s_{new}^2}{n_{new}} \right)^2}{\frac{s_{old}^2}{n_{old}} + \frac{s_{new}^2}{n_{new}}} = \frac{\left(\frac{5.1429}{11} + \frac{6.8571}{10} \right)^2}{\frac{5.1429}{11} + \frac{6.8571}{10}} = \frac{2.2500}{0.1640} = 13.72 = 13$

$|t_{calc}| > |t_{crit}|?$
 $1.6330 > 2.145?$

"The mean recovery times for the old and new procedure are not significantly different ($0.1 < p < 0.2$).

Example #2 - homoscedastic test

Old: $\bar{x}_{old} = 16.00, s_{old}^2 = 5.1429$
New: $\bar{x}_{new} = 14.00, s_{new}^2 = 6.8571$
 $df = 14$ $t_{calc} = 1.6330$

$|t_{calc}| > |t_{crit}|?$
 $1.6330 > 2.145?$

"The mean recovery times for the old and new procedure are not significantly different ($0.1 < p < 0.2$).

Example #2 - homoscedastic test

Old: $\bar{x}_{old} = 16.00, s_{old}^2 = 5.1429$ $df = n_o + n_n - 2 = 8 + 8 - 2 = 14$
New: $\bar{x}_{new} = 14.00, s_{new}^2 = 6.8571$

$s_p^2 = \frac{(n_{old} - 1)s_{old}^2 + (n_{new} - 1)s_{new}^2}{(n_{old} - 1) + (n_{new} - 1)} = \frac{(8 - 1)5.1429 + (8 - 1)6.8571}{(8 - 1) + (8 - 1)} = 6.0000$

$t_{calc} = \frac{\bar{x}_{old} - \bar{x}_{new}}{\sqrt{s_p^2 \left(\frac{1}{n_{old}} + \frac{1}{n_{new}} \right)}} = \frac{16.00 - 14.00}{\sqrt{6.0000 \left(\frac{1}{8} + \frac{1}{8} \right)}} = \frac{2.00}{2.29129} = 1.6330$

$|t_{calc}| > |t_{crit}|?$
 $1.6330 > 2.148?$

"The mean recovery times for the old and new procedure are not significantly different ($0.1 < p < 0.2$).

Example #3 - homoscedastic test

Site 1: $\bar{x}_1 = 11.00, s_1^2 = 23.600$ $df = n_1 + n_2 - 2 = 6 + 8 - 2 = 12$
Site 2: $\bar{x}_2 = 6.00, s_2^2 = 14.000$

$s_p^2 = \frac{(n_1 - 1)s_1^2 + (n_2 - 1)s_2^2}{(n_1 - 1) + (n_2 - 1)} = \frac{(6 - 1)23.600 + (8 - 1)14.000}{(6 - 1) + (8 - 1)} = 18.0000$

$t_{calc} = \frac{\bar{x}_1 - \bar{x}_2}{\sqrt{s_p^2 \left(\frac{1}{n_1} + \frac{1}{n_2} \right)}} = \frac{11.00 - 6.00}{\sqrt{18.0000 \left(\frac{1}{6} + \frac{1}{8} \right)}} = \frac{5.00}{2.29129} = 2.1822$

$|t_{calc}| > |t_{crit}|?$
 $2.1822 > 2.179?$

"The mean number of cracks at site 1 is significantly larger than the mean number of cracks at site 2 ($0.04 < p < 0.05$).

Example #3 - homoscedastic test

Site 1: $\bar{x}_1 = 11.00, s_1^2 = 23.600$
Site 2: $\bar{x}_2 = 6.00, s_2^2 = 14.000$
 $df = 12$ $t_{calc} = 2.1822$

$|t_{calc}| > |t_{crit}|?$
 $2.1822 > 2.179?$

"The mean number of cracks at site 1 is significantly larger than the mean number of cracks at site 2 ($0.04 < p < 0.05$).

Example #3 - heteroscedastic test

Site 1: $\bar{x}_1 = 11.00, s_1^2 = 23.600$
Site 2: $\bar{x}_2 = 6.00, s_2^2 = 14.000$

$t_{calc} = \frac{\bar{x}_1 - \bar{x}_2}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}} = \frac{11.00 - 6.00}{\sqrt{\frac{4.8580}{6} + \frac{3.7417}{8}}} = \frac{5.00}{2.38397} = 2.0973$

$df = \frac{\left(\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2} \right)^2}{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}} = \frac{\left(\frac{4.8580}{6} + \frac{3.7417}{8} \right)^2}{\frac{4.8580}{6} + \frac{3.7417}{8}} = \frac{32.3003}{3.5317} = 9.1458 = 9$

$|t_{calc}| > |t_{crit}|?$
 $2.0973 > 2.262?$

"The mean number of cracks at site 1 is not significantly different from the mean number of cracks at site 2 ($0.05 < p < 0.1$).

Example #3 - heteroscedastic test

Site 1: $\bar{x}_1 = 11.00, s_1^2 = 23.600$
Site 2: $\bar{x}_2 = 6.00, s_2^2 = 14.000$
 $df = 9$ $t_{calc} = 2.0973$

$|t_{calc}| > |t_{crit}|?$
 $2.0973 > 2.262?$

"The mean number of cracks at site 1 is not significantly different from the mean number of cracks at site 2 ($0.05 < p < 0.1$).

Heteroscedastic vs homoscedastic two-sample t-tests

- When the sample sizes are equal, the t_{calc} value will be the same.
- The homoscedastic t-test has slightly more power for 2 reasons:
1. t_{calc} will be equal or larger than heteroscedastic.
2. Degrees of freedom will be larger.
- The homoscedastic t-test does carry an extra risk of type II error when we decide that the population variances are equal.
- Advice: stick to heteroscedastic t-test. Tests only differ if $p < 0.05$, when we should be extra cautious anyway.