

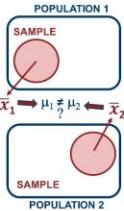
Two sample paired t-test: examples

Review: two-sample 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 are estimates of the population means, but sampling error makes them inexact.

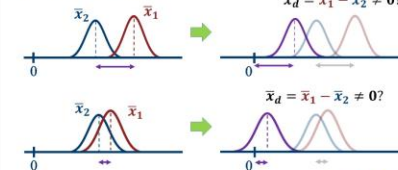
What are the chances the population means are the same (i.e., H_0), based on how much the sample means differ from one another?



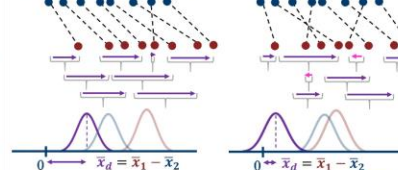
Two-sample t-test: unpaired vs paired data values



Paired t-test concept 1



Two-sample t-test concept 2



Two-sample t-test equation overall

Conduct a one-sample t-test with a hypothetical mean of zero

$$t_{calc} = \frac{\bar{x}_d - 0}{\frac{s_d}{\sqrt{n}}} = \frac{\bar{x}_d}{\frac{s_d}{\sqrt{n}}}$$

$$df = n - 1$$

Example #1 - fin ray symmetry in fish

$\bar{x}_d = -3.000, s_d = 3.6056, n = 9$
 $df = 8$
 $t_{calc} = -2.4962$
 $-2.4962 < -2.306?$

"The mean numbers of fin rays of the left and right sides are significantly different"

Left	Right	Diff.
36	42	-6.00
41	38	3.00
44	43	1.00
40	45	-5.00
39	42	-3.00
37	45	-8.00
30	38	-6.00
41	43	-2.00
34	35	-1.00

Example #1 - fin ray symmetry in fish

$\bar{x}_d = -3.000, s_d = 3.6056, n = 9$
 $t_{calc} = \frac{\bar{x}_d}{\frac{s_d}{\sqrt{n}}} = \frac{-3.000}{\frac{3.6056}{\sqrt{9}}} = -2.4962$
 $df = 9 - 1 = 8$

Pairing: each individual fish creates a distinct left/right pair.



Example #1 - fin ray symmetry in fish

Left and right fin rays of fish. Are the means of the two sides equal? With what confidence do we know?

Second step: means of samples, mean and standard deviation for the differences

Left	Right	Diff.
36	42	-6.00
41	38	3.00
44	43	1.00
40	45	-5.00
39	42	-3.00
37	45	-8.00
30	38	-6.00
41	43	-2.00
34	35	-1.00

Left: $\bar{x}_L = 38.00$
 Right: $\bar{x}_R = 41.00$
 Diff.: $\bar{x}_d = -3.000, s_d = 3.6056, n = 9$



Example #1 - fin ray symmetry in fish

Left and right fin rays of fish. Are the means of the two sides equal? With what confidence do we know?

Left	Right	Diff.
36	42	-6.00
41	38	3.00
44	43	1.00
40	45	-5.00
39	42	-3.00
37	45	-8.00
30	38	-6.00
41	43	-2.00
34	35	-1.00

First step: create a set of the differences.

These are the values we will actually use for the t-test



Example #1 - fin ray symmetry in fish

"The mean numbers of fin rays of the left and right sides are significantly different"

This is missing 2 vital pieces of information.

- Which side has the larger mean, and which has the smaller mean?
 - We have this information, why withhold it?
- What degree of confidence do we have?
 - $p = 0.05$, then the data just barely convinces us.
 - $p < 0.05$, then the data is very convincing.

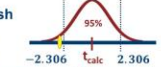


Example #1 - fin ray symmetry in fish

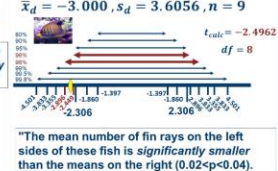
$\bar{x}_d = -3.000, s_d = 3.6056, n = 9$
 $df = 8$
 $t_{calc} = -2.4962$

"The mean numbers of fin rays of the left and right sides are significantly different"

- Which side is larger?
- Using the sign of t_{calc}
 - If t_{calc} is positive, then mean of first data set is larger.
 - If t_{calc} is negative, then mean of second data set is larger.
 - Go back and look at the two sample means to answer this.



Example #1 - fin ray symmetry in fish



"The mean number of fin rays on the left sides of these fish is significantly smaller than the means on the right (0.02 < p < 0.04)." />

spp. A	spp. B	Diff.
22	16	6.00
18	21	-3.00
23	19	4.00
25	18	7.00
25	22	3.00
18	21	-3.00
22	15	7.00
24	22	2.00

Example #2 - flowers in fields

Transects for flowers, count two species. Are the means of the populations equal? With what confidence do we know?

First step: means of samples, mean and standard deviation for the differences

spp. A	spp. B	Diff.
22	16	6.00
18	21	-3.00
23	19	4.00
25	18	7.00
25	22	3.00
18	21	-3.00
22	15	7.00
24	22	2.00

spp. A: $\bar{x}_1 = 22.125$
 spp. B: $\bar{x}_2 = 19.250$
 Diff.: $\bar{x}_d = 2.875, s_d = 4.0510, n = 8$



Example #2 - flowers in fields

$\bar{x}_d = 2.875, s_d = 4.0510, n = 8$
 $t_{calc} = \frac{\bar{x}_d}{\frac{s_d}{\sqrt{n}}} = \frac{2.875}{\frac{4.0510}{\sqrt{8}}} = 2.0073$
 $df = 8 - 1 = 7$

Pairing: each transect creates a distinct pair of values.

Paired vs unpaired two-sample t-tests

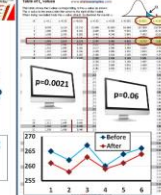
- Pairing must be based on inherent criteria, we can't just choose the pairs.
 - Transects, left vs right, before vs after, twins, clones, etc.
- Even if a paired design goes bad, the unpaired test could be used.
- If possible, use the paired test since it is much more powerful.

Example #3 - Cholesterol

$\bar{x}_d = 3.167, s_d = 1.4720, n = 6$
 $df = 5$
 $t_{calc} = 5.2697$
 $5.2697 > 2.571?$

"The mean cholesterol value after treatment is significantly lower than the mean value before treatment (0.002 < p < 0.005)"

Significant ✓ Relevant ?



Example #3 - Cholesterol

$\bar{x}_d = 3.167, s_d = 1.4720, n = 6$
 $t_{calc} = \frac{\bar{x}_d}{\frac{s_d}{\sqrt{n}}} = \frac{3.167}{\frac{1.4720}{\sqrt{6}}} = 5.2697$
 $df = 6 - 1 = 5$

Pairing: each subject creates a distinct pair of before/after values.



Example #3 - Cholesterol

Test a new drug, measure before & after. Are the means before & after equal? With what confidence do we know?

First step: means of samples, mean and standard deviation for the differences

Before	After	Diff.
265	261	4.00
262	257	5.00
267	263	4.00
260	259	1.00
264	261	3.00
266	264	2.00

Before: $\bar{x}_{before} = 264.000$
 After: $\bar{x}_{after} = 260.833$
 Diff.: $\bar{x}_d = 3.167, s_d = 1.4720, n = 6$



Example #2 - flowers in fields

$\bar{x}_d = 2.875, s_d = 4.0510, n = 8$
 $df = 7$
 $t_{calc} = 2.0073$
 $2.0073 > 2.365?$

"The mean numbers of each flower species are not significantly different (0.05 < p < 0.1)"

p = 0.0847

