

⚠ Lagunita is retiring and will shut down at 12 noon Pacific Time on March 31, 2020. A few courses may be open for self-enrollment for a limited time. We will continue to offer courses on other online learning platforms; visit <http://online.stanford.edu>.

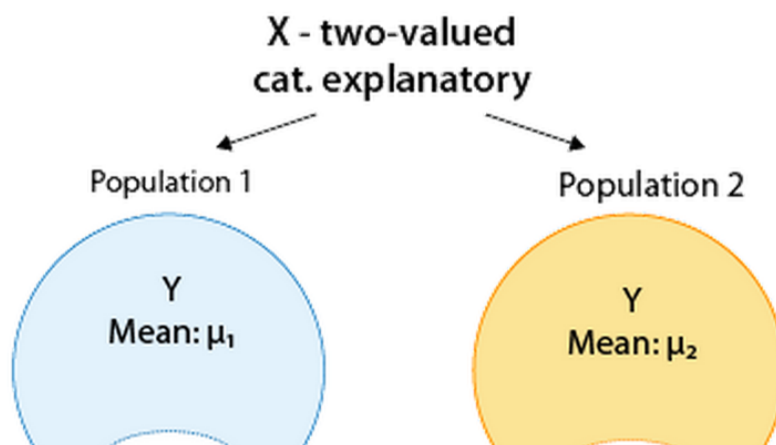
Course > Inference: Relationships C→Q > Matched Pairs > Matched Pairs: Summary

🔖 Bookmark this page

Matched Pairs: Summary

Let's Summarize

- The paired t-test is used to compare two population means when the two samples (drawn from the two populations) are dependent in the sense that every observation in one sample can be linked to an observation in the other sample. Such a design is called "matched pairs."



- The most common case in which the matched pairs design is used is when the same subjects are measured twice, usually before and then after some kind of treatment and/or intervention. Another classic case are studies involving twins.
- The idea behind the paired t-test is to reduce the data from two samples to just one sample of the differences, and use these observed differences as data for inference about a single mean—the mean of the differences, μ_d .

Pairs	1	2	3	4	...	n
Sample 1	*	*	*	*	...	*
Sample 2	*	*	*	*	...	*
Differences sample1-sample2	d_1	d_2	d_3	d_4	...	d_n

- The paired t-test is therefore simply a one-sample t-test for the mean of the differences μ_d , where the null value is 0. The null hypothesis is therefore:

- $H_0: \mu_d = 0$

and the alternative hypothesis is one of the following (depending on the context of the problem):

- $H_a: \mu_d < 0$
- $H_a: \mu_d > 0$
- $H_a: \mu_d \neq 0$

- The paired t-test can be safely used when one of the following two conditions hold:
 - The differences have a normal distribution.
 - The sample size of differences is large.

When the sample size of difference is not large (and we therefore need to check the normality of the differences), what we do in practice is look at the histograms of the differences and make sure that there are no signs of non-normality like extreme skewedness and/or outliers.

- The test statistic is as follows and has a t distribution when the null hypothesis is true:

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

- P-values are obtained from the output, and conclusions are drawn as usual, comparing the p-value to the significance level α .
- If H_0 is rejected, a 95% confidence interval for μ_d can be very insightful and can also be used for the two-sided test.

Open Learning Initiative [🔗](#)



[🔗](#) Unless otherwise noted this work is licensed under a Creative Commons Attribution-NonCommercial-ShareAlike 4.0 International License [🔗](#).

© All Rights Reserved