Two Sample *t* Tests for Comparing Difference between Means

Sunthud Pornprasertmanit Chulalongkorn University

Statistical Inference

If you have the parameters, you must not analyze any statistical inference.

You conduct statistical inference only in

1) Predictions about populations whose element are so numerous that viewing them all is impossible
2) Predictions about phenomena that cannot be directly observed

Dependent and Independent Groups

In comparing sample means, you must consider research question and design.

For example

1) Do women have more self-discipline than men?
2) Is psychotherapy program effective in treating depressed people?
3) Is mnemonics method more effective than normal one?
4) Is child-center teaching more effective than lecture-based teaching in computer study when control for English proficiency?

Samples are independent if the selection of elements in one sample is not affected by the selection of elements in the other. There are two designs related to independent samples.

1) Random samples from different population
2) Random assignment of units into different groups

Samples are dependent if the selection of elements in one sample is affected by the selection of elements in the other.

1) Repeated measures in the same unit.
2) Matching design or randomized block design
3) Twin matching
4) Matching by mutual selection

Author Note
This article was written in September 2007 for teaching in Introduction to Statistics in Psychology Class, Faculty of Psychology, Chulalongkorn University
Correspondence to Sunthud Pornprasertmanit. Email: psunthud@gmail.com

Independent *t* Test for Difference between Means

Although the population difference is equal to zero, the differences between two sample means are not equal to zero by chance.

Therefore, the difference between two sample means must be proved that this difference did not occur by sampling error.

Null hypothesis $H_0: \mu_1 - \mu_2 = \delta_0; \delta = \delta_0$

Alternative hypothesis $H_1: \mu_1 - \mu_2 \neq \delta_0; \delta \neq \delta_0$ (Two-tailed)

$H_1: \mu_1 - \mu_2 > \delta_0; \delta > \delta_0$ (One-tailed)

$H_1: \mu_1 - \mu_2 < \delta_0; \delta < \delta_0$ (One-tailed)

From general form of *t* statistics,

$$t = \frac{\text{Statistics} - \text{Null hypothetical value}}{\text{Estimated Standard Error of Statistics}}$$

Statistics is the difference between sample means.

Null hypothetical value is 0 or specified value.

Estimated standard error of statistics is

$$Var(\bar{X}_1 - \bar{X}_2) = \hat{\sigma}^2_{Pooled}\left(\frac{1}{n_1} + \frac{1}{n_2}\right)$$

$$\hat{\sigma}^2_{Pooled} = \frac{(n_1 - 1)\hat{\sigma}^2_1 + (n_2 - 1)\hat{\sigma}^2_2}{(n_1 - 1) + (n_2 - 1)}$$

Therefore,

$$t = \frac{\hat{\delta} - \delta_0}{\sqrt{\hat{\sigma}^2_{Pooled}\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}}$$

If null hypothesis is $H_0: \mu_1 - \mu_2 = 0$, the formula is

$$t = \frac{\bar{X}_1 - \bar{X}_2}{\sqrt{\hat{\sigma}^2_{Pooled}\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}}$$

Degree of freedom is equal to $(n_1 - 1) + (n_2 - 1) = n_1 + n_2 - 2$

The confidence interval based on this null hypothesis testing is

A two-sided 100(1-α) % confidence interval for $\mu_1 - \mu_2$ is given by

$$\bar{X}_1 - \bar{X}_2 - t_{\alpha/2,df}\sqrt{\hat{\sigma}^2_{Pooled}\left(\frac{1}{n_1}+\frac{1}{n_2}\right)} < \mu_1 - \mu_2 < \bar{X}_1 - \bar{X}_2 + t_{\alpha/2,df}\sqrt{\hat{\sigma}^2_{Pooled}\left(\frac{1}{n_1}+\frac{1}{n_2}\right)}$$

Lower and upper one-sided 100(1-α) % confidence intervals for $\mu_1 - \mu_2$ are given by

$$\bar{X}_1 - \bar{X}_2 - t_{\alpha,df}\sqrt{\hat{\sigma}^2_{Pooled}\left(\frac{1}{n_1}+\frac{1}{n_2}\right)} < \mu_1 - \mu_2$$

and $$\mu_1 - \mu_2 < \bar{X}_1 - \bar{X}_2 + t_{\alpha,df}\sqrt{\hat{\sigma}^2_{Pooled}\left(\frac{1}{n_1}+\frac{1}{n_2}\right)}$$

Example

Assumption of this null hypothesis testing and confidence interval

1) Two random samples from two population or randomly assigned to two groups
2) Two populations are normally distributed
3) The variance of the populations are unknown but are assumed to be equal (homogeneity of variance)

The *t* statistic is robust with respect to violation of normality assumption if two sample sizes are equal and the two population distributions have similar shapes, unimodal, and no outliers.

When n_1 and n_2 both greater than 30, the normality assumption is no longer important because of the central limit theorem.

The two-sample *t* test for independent samples is robust with respect to violation of the assumption of equal population variances, provided that $n_1 = n_2$.

If the population variances are unequal and the sample *n*'s are unequal, the sample variances should not be pooled. The modified statistic in next section can be used.

Levene's test can be used for measuring the violation of homogeneity of variance. (Explain after studying one-way ANOVA)

Independent *t* Test for Difference between Means when Equal Variances not Assumed

The Welch Procedure

The modified statistic is

$$t = \frac{\hat{\delta} - \delta_0}{\sqrt{\hat{\sigma}_1^2/n_1 + \hat{\sigma}_2^2/n_2}}$$

Degree of freedom of this test is equal to

$$v' = \frac{\left(\frac{\hat{\sigma}_1^2}{n_1} + \frac{\hat{\sigma}_2^2}{n_2}\right)^2}{\frac{1}{n_1 - 1}\left(\frac{\hat{\sigma}_1^2}{n_1}\right)^2 + \frac{1}{n_2 - 1}\left(\frac{\hat{\sigma}_2^2}{n_2}\right)^2}$$

The confidence interval based on this null hypothesis testing is

A two-sided 100(1-α) % confidence interval for $\mu_1 - \mu_2$ is given by

$$\bar{X}_1 - \bar{X}_2 - t_{\alpha/2,v'}\sqrt{\frac{\hat{\sigma}_1^2}{n_1} + \frac{\hat{\sigma}_2^2}{n_2}} < \mu_1 - \mu_2 < \bar{X}_1 - \bar{X}_2 + t_{\alpha/2,v'}\sqrt{\frac{\hat{\sigma}_1^2}{n_1} + \frac{\hat{\sigma}_2^2}{n_2}}$$

Lower and upper one-sided 100(1-α) % confidence intervals for $\mu_1 - \mu_2$ are given by

$$\bar{X}_1 - \bar{X}_2 - t_{\alpha,v'}\sqrt{\frac{\hat{\sigma}_1^2}{n_1} + \frac{\hat{\sigma}_2^2}{n_2}} < \mu_1 - \mu_2 \qquad \text{and} \qquad \mu_1 - \mu_2 < \bar{X}_1 - \bar{X}_2 + t_{\alpha,v'}\sqrt{\frac{\hat{\sigma}_1^2}{n_1} + \frac{\hat{\sigma}_2^2}{n_2}}$$

Example

Assumption of this null hypothesis testing and confidence interval

1) Two random samples from two population or randomly assigned to two groups
2) Two populations are normally distributed
3) Both sample sizes are 5 or more.

Independent *z* Test for Difference between Means when Population Variances are known

When the population variances are known, researchers may use *z* test for testing differences between samples.

$$z = \frac{\hat{\delta} - \delta_0}{\sqrt{\sigma_1^2/n_1 + \sigma_2^2/n_2}}$$

However, the population variances are rarely known. Then, this statistic is rarely used.

Practical Significance of Independent *t* Test for Difference between Means

Hedges' g statistic

$$g = \frac{|\bar{X}_1 - \bar{X}_2|}{\hat{\sigma}_{Pooled}}$$

$$\hat{\sigma}_{Pooled} = \sqrt{\frac{(n_1 - 1)\hat{\sigma}_1^2 + (n_2 - 1)\hat{\sigma}_2^2}{(n_1 - 1) + (n_2 - 1)}}$$

If the assumption that the population variances are equal is not tenable, the variances should not be pooled in computing g. Sample standard deviation of control group or the standard deviation of group that is used as baseline is used in place of pooled standard deviation.

The Hedges' *g* statistic formula when know *t* statistic

If two group sample sizes are equal,

$$g = \frac{\sqrt{2}|t|}{\sqrt{n}}$$

If two group sample sizes are not equal

$$g = \frac{|t|\sqrt{n_1 + n_2}}{\sqrt{n_1 n_2}}$$

Dependent *t* Test for Difference between Means

The dependent t test is used when the observations in both groups are not statistical independent. (The four conditions in the beginning of this lecture)

Two formula of dependent *t* test

Null hypothesis	$H_0: \mu_1 - \mu_2 = \delta_0; \delta = \delta_0$	
Alternative hypothesis	$H_1: \mu_1 - \mu_2 \neq \delta_0; \delta \neq \delta_0$	(Two-tailed)
	$H_1: \mu_1 - \mu_2 > \delta_0; \delta > \delta_0$	(One-tailed)
	$H_1: \mu_1 - \mu_2 < \delta_0; \delta < \delta_0$	(One-tailed)

The formula of dependent *t* test is

$$t = \frac{\bar{X}_1 - \bar{X}_2}{\sqrt{\frac{\hat{\sigma}^2_{Pooled}}{n_1} + \frac{\hat{\sigma}^2_{Pooled}}{n_1} - 2r_{12}\left(\frac{\hat{\sigma}^2_{Pooled}}{\sqrt{n_1}}\right)\left(\frac{\hat{\sigma}^2_{Pooled}}{\sqrt{n_2}}\right)}}$$

If the correlation between two groups increases, the standard error decreases and then the *t* statistic increase.

The degree of freedom of this statistic is $n - 1$, if n is the number of pairs.

However, a simpler alternative formula is available.

$$D_i = X_{i1} - X_{i2}$$

$$t = \frac{\bar{X}_D - \delta_0}{\hat{\sigma}_{\bar{X}_D}} = \frac{\bar{X}_D - \delta_0}{\frac{\hat{\sigma}_D}{\sqrt{n}}}$$

The degree of freedom of this statistic is n – 1, if n is the number of difference scores.

This formula likes one-sample *t* test.

The confidence interval based on this null hypothesis testing is

A two-sided 100(1-α) % confidence interval for $\mu_1 - \mu_2$ is given by

$$\bar{X}_D - t_{\alpha/2,df}\hat{\sigma}_{\bar{X}_D} < \mu_1 - \mu_2 < \bar{X}_D + t_{\alpha/2,df}\hat{\sigma}_{\bar{X}_D}$$

Lower and upper one-sided 100(1-α) % confidence intervals for $\mu_1 - \mu_2$ are given by

$$\bar{X}_D - t_{\alpha,df}\hat{\sigma}_{\bar{X}_D} < \mu_1 - \mu_2 \qquad \text{and} \qquad \mu_1 - \mu_2 < \bar{X}_D + t_{\alpha,df}\hat{\sigma}_{\bar{X}_D}$$

Example

Assumption of this null hypothesis testing and confidence interval

1) The pairs of data are randomly drawn from population.
2) The population of differences is normally distributed. These differences will be normally distributed if X_1 and X_2 are normally distributed.
3) The standard error of the mean of the difference scores is unknown and must be estimated from sample data.

Practical Significance of Dependent *t* Test for Difference between Means

Hedges' g statistic

$$g = \frac{|\bar{X}_1 - \bar{X}_2|}{\hat{\sigma}_D}$$

The Hedges' *g* statistic formula when know *t* statistic

$$g = |t|\sqrt{\frac{2\hat{\sigma}_D^2}{n(\hat{\sigma}_1^2 + \hat{\sigma}_2^2)}}$$

Sample Size of Independent *t* Test for Difference between Means

Appendix D8 in Kirk (2008) can be used to select sample sizes for two-sample *t*-test

To estimate required sample sizes for independent t test, it is necessary to specify α, β, and Cohen's *d*, but, for dependent *t* test, the correlation between the two populations is required for estimating sample size.

The correlation between the two populations is rarely known, its estimation must be based on previous research or informed judgment. If researchers are not confident of their estimates of ρ, they can use a conservative estimate.

Casual Relationship

The conditions that make causal relation are

1) X precedes Y in time
2) Some mechanism explained
3) Change in X is accompanied by change in Y
4) Effect X on Y cannot be explained by other variables

One of the types of research that prove casual relation is experimental design.

Experimental Design

Experimental studies have at least three special characteristics

1) Random assignment of cases to levels of IVs
2) Manipulation of those levels
3) Control of extraneous variables

Random assignment is used for at least two reasons

1) To remove researcher bias from the study – to keep the researcher from assigning (however unconsciously) bigger, stronger, smarter to the level of treatment he or she hopes is more effective
2) To let random processes average out differences among cases

Manipulation of the levels of treatment is undertaken so that the researcher controls both the precise nature of each level of treatment and the timing of its delivery.

Control of extraneous variables ensures that the only thing that changes during the study is level of treatment. The ways to control extraneous variables are

1) Hold extraneous variables constant
2) Random assignment
3) Matching or randomized block design
4) Statistical control or turn extraneous variable into another IV

The extraneous variables can occur before or during experiment.

Quasi-experimental research is the experimental research that cannot randomly assign cases to levels of IV.

The quasi-experimental research may conclude casual relation if designing with carefulness.

In conclusion, the statistical analysis cannot conclude casual relation, but good experimental design and good statistical analysis can conclude casual relation.