Testing Differences between Two Means

Large Independent Sample Means:

Used to test whether the data from two samples come from the same populations or whether two populations are different.

Assumptions:

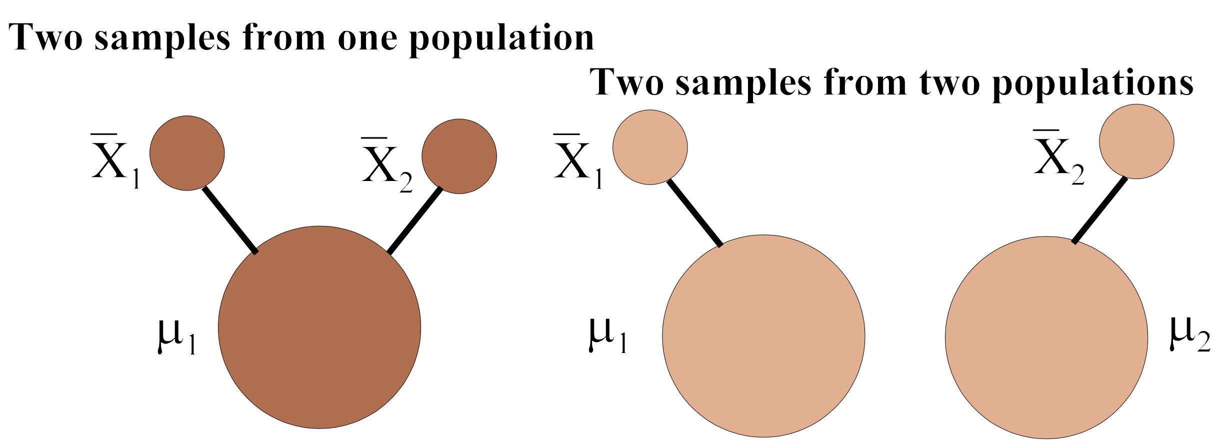
- samples must be independent, i.e., there can be no relationship between the two samples
- populations must be normally distributed and standard deviations known or sample size > 30
- should not be used if more than two means are tested unless adjustments are made to significance levels (e.g., Bonferroni correction, $\alpha_{Bonferroni} = \alpha$/number of tests)

Z-test:

Test value:

$$z = \frac{(\overline{X}_1 - \overline{X}_2) - (\mu_1 - \mu_2)}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}}$$

Critical value comes from standard normal (*z*) distribution. Use one- or two-tailed test. Conservatively, choose the two-tailed test. Values are also available at bottom of *t*-distribution.



The Step-by-Step Approach

Step 1: State hypotheses

Two-tailed:

$H_0: \mu_1 = \mu_2$

$H_1: \mu_1 \neq \mu_2$

One-tailed:

$H_0: \mu_1 \leq \mu_2$ or $H_0: \mu_1 \geq \mu_2$

$H_1: \mu_1 > \mu_2$ or $H_1: \mu_1 < \mu_2$

Step 2: Find critical value

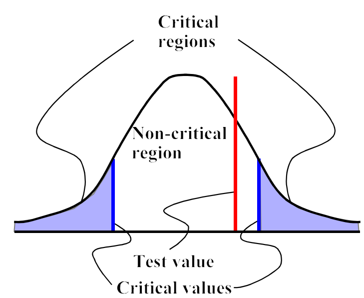
Look up z-score for specified significance (α) level and for one- or two-tailed test (selected in advance). Usually use $\alpha = 0.05$ and two-tailed test, i.e., $z_{critical} = \pm 1.960$. For one-tailed use $z_{critical} = \pm 1.645$.

Step 3: Compute test value

$$z = \frac{(\overline{X}_1 - \overline{X}_2) - (\mu_1 - \mu_2)}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}}$$

Step 4: Make decision

Draw diagram of normal distribution and critical regions. If test value is in critical region reject the null hypothesis otherwise do not reject.



Step 5: Summarize results

Restate hypothesis (null or alternate) accepted in step 4.

If reject null:

There is enough evidence to reject the null hypothesis.

If not reject null:

There is not enough evidence to reject the null hypothesis.

Optionally, reword hypothesis in "lay" terms. E.g., There is/is not a difference between the two populations **or** one population is greater/lesser than the other for the independent variable.

Testing Differences between Two Means

Small Independent Sample Means:

When population standard deviations are unknown **and** sample size is < 30 use the t-distribution for critical values and a t-test for test values. First use an F-ratio to determine whether sample variances are equal or unequal. Then choose the correct t-test.

Assumptions

- two samples must be independent, i.e., different subjects—if not, use "dependent-groups t-test"
- data must be normally distributed

If sample variances are NOT equal:

Use test value:

$$t = \frac{(\overline{X}_1 - \overline{X}_2) - (\mu_1 - \mu_2)}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}}$$

For degrees of freedom (df) use smaller of $n_1 - 1$ and $n_2 - 1$
(i.e., conservative choice, higher critical value)

If sample variances are equal:

Use test value:

$$t = \frac{(\overline{X}_1 - \overline{X}_2) - (\mu_1 - \mu_2)}{\sqrt{\frac{(n_1 - 1)s_1^2 + (n_2 - 1)s_2^2}{n_1 + n_2 - 2}}\sqrt{\frac{1}{n_1} + \frac{1}{n_2}}}$$

and $df = n_1 + n_2 - 2$

Uses a "pooled" estimate of variance that combined with a larger degree of freedom (n_1+n_2-2) increases the test's power (i.e., ability to find a true difference).

Test for Equal Variances

Also called Homogeneity of Variance

- used primarily to determine which *t*-test to use
- uses *F*-distribution and *F*-test (later used for ANOVA)
- assume variances are equal and test if unequal
- SPSS uses "Levine's Test for Equality of Variances"
  If *P* (Sig.) < α variances are NOT equal.

Step 1: Always a two-tailed test.

$$\mathbf{H_0:}\ s_1^2 = s_2^2$$

$$\mathbf{H_1:}\ s_1^2 \neq s_2^2$$

Step 2:

Find critical value (F_{CV}) from F-distribution. Use degrees of freedom of larger variance ($df_N = n_{\text{larger}} - 1$) as numerator and degrees of freedom of smaller variance as denominator ($df_D = n_{\text{smaller}} - 1$).

Step 3:

Compute test value: $$F_{TV} = \frac{s^2_{\text{larger}}}{s^2_{\text{smaller}}}$$

Note, F_{TV} will always be ≥ 1.

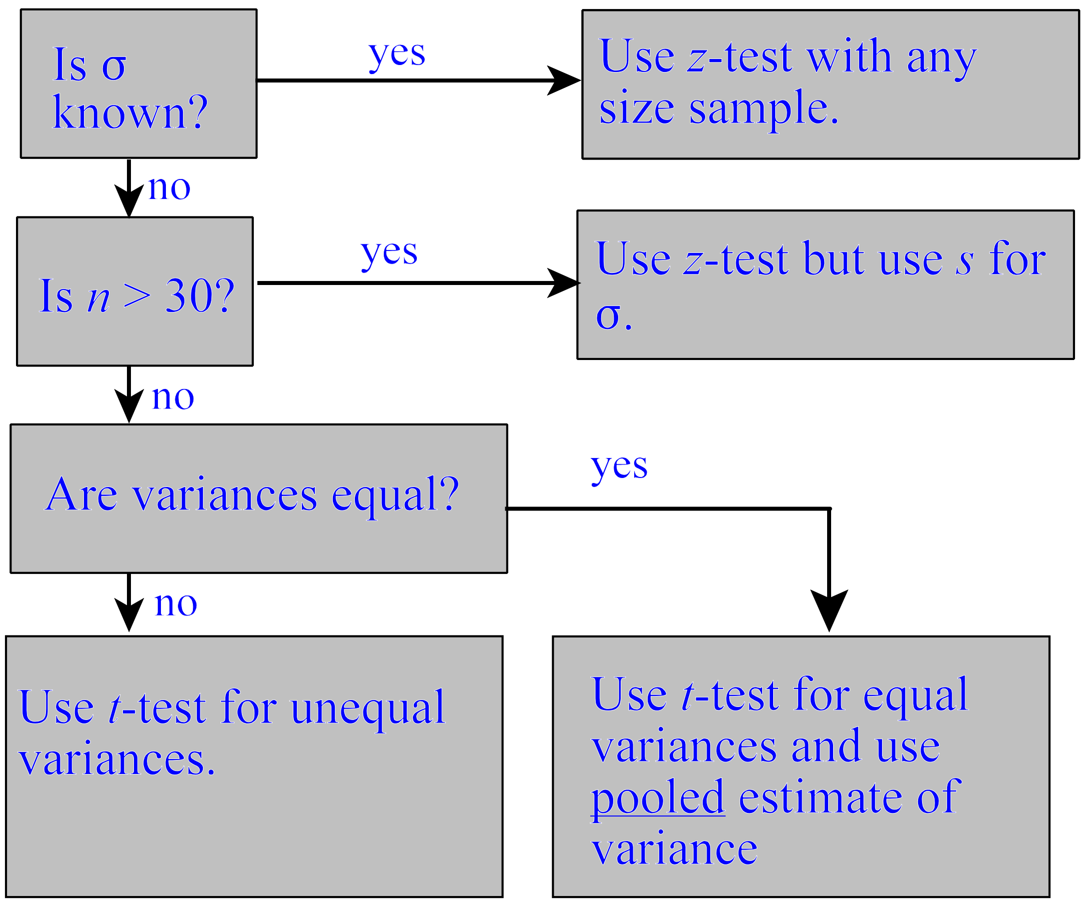
Step 4 and 5:

If $\boldsymbol{F_{TV}} > \boldsymbol{F_{CV}}$ then reject H_0 and conclude variances are unequal.

If $\boldsymbol{F_{TV}} \leq \boldsymbol{F_{CV}}$ then do NOT reject H_0 and conclude variances are equal. I.e., you have **homogeneity of variances**. You can now select the appropriate "Independent-groups *t*-test".

Flow Diagram for Choosing the Correct Independent Samples *t*-Test

Similar to flow diagram used for single sample means. But requires a test for equality of variances **(homogeneity of variance**). Generally the sample's mean and standard deviation are used with the *t*-distribution. The *t*-distribution becomes indistinguishable from the *z*-distribution (normal distribution) when $n > 30$. Samples must be independent.



Testing Differences between Two Means

Dependent Sample Means:

Used when two samples are not independent. More powerful then independent groups *t*-test and easier to perform (no variance test required). Simplifies research protocol (i.e., fewer subjects) but dependence may limit generalizability.

Examples:

- repeated measures (test/retest, before/after)
- matched pairs *t*-test (subjects matched by a relevant variable: height, weight, shoe size, IQ score, age)
- twin studies (identical, heterozygotic, living apart)

Step 1:

Two-tailed:	One-tailed:		
$H_0: \mu_D = 0$	$H_0: \mu_D \leq 0$	or	$H_0: \mu_D \geq 0$
$H_1: \mu_D \neq 0$	$H_1: \mu_D > 0$	or	$H_1: \mu_D < 0$

Step 2:

Critical value from ***t*-distribution** with degrees of freedom equal to number of data pairs minus one (***df* = *n* – 1**).

Step 3:

Compute differences between pairs (**D**) then mean difference ($\overline{D}$) and s_D:

$$\overline{D} = \frac{\sum D}{n} \quad \text{and} \quad s_D = \sqrt{\frac{\Sigma D^2 - \frac{(\Sigma D)^2}{n}}{n - 1}}$$

Test value:

$$t = \frac{\overline{D} - \mu_D}{\frac{s_D}{\sqrt{n}}}$$

Step 4 and 5:

If test value > critical value reject H_0 otherwise there is no difference between the two trials/groups.