September 27, 2005

# Documentation for Two-sample Independent *t* Test


Minn M. Soe, MD, MPH, MCTM : msoe@sph.emory.edu
Kevin M. Sullivan, PhD, MPH, MHA: cdckms@sph.emory.edu


This analysis is conducted to observe whether there is a significant difference in means between two independent samples, given respective standard deviation or standard error. The data input screen is as follows:

Calculate

Clear

| Two-Sample Independent *t* Test | | | | | |
|---|---|---|---|---|---|
| Confidence Interval (%) {two-sided} | | | 95 | *Enter a value between 0 and usually 95%* | |
| | Sample Size | Mean | Std. Dev. | (or) | Std. E |
| Group 1 | 7 | 11.57 | 8.81 | | |
| Group 2 | 18 | 7.44 | 3.698 | | |

The input values requested are:

- Two-sided confidence intervals (%) that can be chosen are between 0 and 100.
- Enter individual sample means.
- Enter standard deviation (or) standard error of individual sample mean.

The result of the calculation is shown next:

## Two-Sample Independent *t* Test

### Input Data

| Two-sided confidence interval | 95% | | | |
|---|---|---|---|---|
| | **Sample size** | **Mean** | **Std. Dev.** | **Std. Error** |
| Group-1 | 7 | 11.57 | 8.81 | |
| Group-2 | 18 | 7.44 | 3.698 | |

| **Result** | ***t* statistics** | ***df*** | **p-value[1]** | **Mean Difference** | **Lower Limit** | **Upper Li** |
|---|---|---|---|---|---|---|
| **Equal variance** | 1.68286 | 23 | 0.105931 | 4.13 | -0.946799 | 9.206 |
| **Unequal variance** | 1.19986 | 7 | 0.269221 | 4.13 | -4.00922 | 12.269 |

| | ***F* statistics** | ***df*(numerator,denominator)** | **p-value[1]** |
|---|---|---|---|
| **Test for equality of variance[2]** | 5.67568 | 6,17 | 0.00429641 |

[1] p-value (two-tailed)

[2] Hartley's *f* test for equality of variance

**Results from OpenEpi open source calculator--t-testMean**

The interpretation of the test is that there is no significant difference between the means of these two groups. It is noteworthy that before interpreting as above, *F* test for the equality of variances from these two independent, normally distributed samples should be first checked. If the two variances are not significantly different, ie. p-value of test for equality of variance is >0.05, the result of difference in means should be interpreted from t statistics and p-value based on equal variance. In the example above, the two variances are significantly different, ie. p-value of *F* test is 0.004, and therefore, the p-value of difference in means is 0.2424.

In addition, the confidence interval of difference in means is also displayed.

**The formulae for two-sample *t* test are as follows:**

All statistics are derived from formulae of the text 'Fundamentals of Biostatistics' ($5^{th}$ edition) by Bernard Rosner; For two-sample t test with equal variance, statistics were based on equation 8.11 to 8.13; If assuming unequal variance, statistics were based on equation 8.21 to 8.23.

- **Two-sample *t* test with equal variance:**

$$t = \frac{[X_1 - X_2]}{S\sqrt{\frac{1}{n_1} + \frac{1}{n_2}}}$$

$$where \quad S = \frac{[(n_1 - 1) \times S_1^2 + (n_2 - 1) \times S_2^2]}{n_1 + n_2 - 2}$$

$$df = (n_1 - 1) + (n_2 - 1) = n_1 + n_2 - 2$$

$$X_1 - X_2 \pm t_{(n_1 + n_2 - 2,\ 1 - \alpha/2)} \times S\sqrt{\frac{1}{n_1} + \frac{1}{n_2}}$$

S = pooled estimate of the variance.
$df$ = degree of freedom

- **Two-sample *t* test with unequal variance:**
  **Satterthwaite's method (see also Welch's modified t test)**

$$t = \frac{[X_1 - X_2]}{\sqrt{\frac{S_1^2}{n_1} + \frac{S_2^2}{n_2}}}$$

$$df = round \quad \frac{\left(\frac{S_1^2}{n_1} + \frac{S_2^2}{n_2}\right)^2}{\frac{(S_1^2 / n_1)^2}{n_1 - 1} + \frac{(S_2^2 / n_2)^2}{n_2 - 1}}$$

$$X_1 - X_2 \pm t_{(df,\ 1-\alpha/2)} \times \sqrt{\frac{S_1^2}{n_1} + \frac{S_2^2}{n_2}}$$

$df$ = approximate degree of freedom.
Standard error=Standard deviation/√n

- **Hartley's *f* test for equal variance:**

$$\boxed{f \text{ statistics} = S^2_L / S^2_S}$$

$S^2_L$ = the larger of two variances;
$S^2_S$ = the smaller of two variances

Note: test for equality of variance is based on equation 8.15-8.16.

**Reference:**


- Bernard Rosner. Fundamentals of Biostatistics (5th edition).
- Welch, B. L. (1938). The significance of the difference between two means when the population variances are unequal. *Biometrika* 29, 350-362.
- Statterthwaite, F. E. (1946). An approximate distribution of estimates of variance components. *Biometrics Bulletin* 2, 110-114.



**Acknowledgement:**
Default values were obtained from example 8.18 (pg. 297-8) described in 'Fundamentals of Biostatistics' (5th edition) by Bernard Rosner.