# Student t-test Summary:

Two Student t-tests show which compounds have the power to differentiate the different two-groups in the data set. The two sample t-test is applied to one metabolite at a time (i.e., a univariate analysis) to determine whether the mean values of the two groups are different. The null hypothesis for the test is H0 : mu\_group1 = mu\_group2, and the alternative hypothesis is Ha : mu\_group1 != mu\_group2. If the p-value for the test is smaller than a cutoff value, typically 0.05, the null hypothesis is rejected. If the p-value is large, there is no significant difference between the mean values for the two groups, indicating the metabolite has little power to separate them.

The 0.05 cutoff value is often used when the t-test for a metabolite is examined individually, without considering the tests for other metabolites. A multiple comparison procedure can be employed, in which a smaller cutoff value is used, to control the overall error caused by using all the t-tests together.

### Input Summary:

Input Dataset: ->e.csv;

Output Dataset: ->Student t-test->student\_t\_test.csv;

**- Treatment Group:** Genotype. The student t-test will be performed on each compound to detect those significantly altered by Genotype.

**- Variance Equality Assumption:** FALSE. If TRUE, the equality of variance is assumed, and the tests are Student's t-test, otherwise Welch t-test.

**- Correct the False Discovery Rate:** Benjamini & Hochberg (1995). When conducting multiple tests, the rate of incorrectly reject a null hypothesis will increase. FDR-controlling procedures are designed to control the expected proportion of "discoveries" (rejected null hypotheses) that are false (incorrect rejections). The suggested method for metabolomics is the Benjamini & Hochberg procedure. For more information, please visit https://en.wikipedia.org/wiki/False\_discovery\_rate.

### Result Summary:

Welch t-test (not assuming equal variance) was performed on each compound to test if the mean average of Lmbrd1 different from Null. Out of 740 compounds, 71 are significant with p-value < 0.05. To control the false discovery rate (FDR), the Benjamini & Hochberg (1995) procedure was used. After FDR correction, 1 compounds are still significant. See Table 1 for more detail.

Table Explanation.

- index: the index of compounds, mainly for sorting the table.

- label: compound labels.

- p\_values: p-values from t-tests.

- p\_values\_adjusted: p-values adjusted by the FDR correction procedure.

| label | p\_values | p\_values\_adjusted |
| --- | --- | --- |
| 3,6-anhydro-D-galactose | 1.245943e-08 | 9.219976e-06 |
| 9,12-Octadecadiynoic Acid | 6.707096e-04 | 2.481625e-01 |
| gamma-Glutamylmethionine | 2.858643e-03 | 4.747872e-01 |
| .gamma.-L-Glu-.epsilon.-L-Lys | 3.590251e-03 | 4.747872e-01 |
| Thiazolidine-4-carboxylic acid | 3.614101e-03 | 4.747872e-01 |
| Bicyclo[2.2.1]heptane-2-methanol | 3.849626e-03 | 4.747872e-01 |
| Palmitamide | 5.212657e-03 | 4.980442e-01 |
| glycerol | 6.045576e-03 | 4.980442e-01 |
| 3-Ureidopropionic acid | 7.647428e-03 | 4.980442e-01 |
| TG (56:7) | 8.109736e-03 | 4.980442e-01 |

Table 1: most significant compounds (i.e. small p-values)