

Conducting a 2-sample t-test using RStudio

These instructions assume that you've already installed the necessary package and that you've imported and attached your data file.

To conduct a 2-sample t-test in Rstudio, you will use the following R command:

`t.test(variable~ group, data=dataset, var.equal=TRUE, paired=FALSE)`

- *variable* = The name of the variable you are exploring.
- *group* = The name of the categorical variable that arranges the data.
- *dataset* = The name of your dataset or spreadsheet.
- *var.equal* = When we perform a 2-sample t-test, we will assume that the variance is equal so this will be set to TRUE. If you know the variance is not equal, then set to FALSE.
- *paired* = Set to FALSE when you know you are not working with paired data.

Let's say you have the following data regarding the test scores from two classes:

Class	TestGrades
Dr. Howard	99
Dr. Howard	93
Dr. Howard	81
Dr. Howard	92
Dr. Howard	89
Dr. Howard	80
Dr. Howard	88
Dr. Howard	81
Dr. Howard	100
Dr. Howard	82
Dr. Smith	98
Dr. Smith	66
Dr. Smith	86
Dr. Smith	62
Dr. Smith	78
Dr. Smith	87
Dr. Smith	77
Dr. Smith	60
Dr. Smith	60
Dr. Smith	79

We want to determine whether or not there is a statistical difference in the mean scores for the two classes. The hypotheses would be as follows:

Ho: The mean of Dr. Howard's class does not differ from the mean of Dr. Smith's class.

Ha: The mean of Dr. Howard's class differs from the mean of Dr. Smith's class.

Symbolically:

$$H_0: \mu_1 = \mu_2$$

$$H_a: \mu_1 \neq \mu_2$$

Here's the command that you would input into R:

```
t.test(TestGrades ~ Class, data=t_test_in_class_dataset, var.equal=TRUE, paired=FALSE)
```

Here's the output:

Two Sample t-test

data: TestGrades by Class

t = 2.7831, df = 18, p-value = 0.01227

alternative hypothesis: true difference in means between group Dr. Howard and group Dr. Smith is not equal to 0

95 percent confidence interval:

3.235436 23.164564

sample estimates:

mean in group Dr. Howard mean in group Dr. Smith

88.5

75.3

The p-value returned was 0.01227 which is less than our alpha of .05. As such, we will reject the null hypothesis and accept the alternative hypothesis that there is difference in the mean scores of the class.

Please note the wording associated with the alternative hypothesis statement in the output. Another way to represent the hypotheses for a 2-sample t-test is as follows:

$$H_0: \mu_1 - \mu_2 = 0$$

$$H_a: \mu_1 - \mu_2 \neq 0$$