Rapport package team

F test

2011-04-26 20:25 CET

## Description

This template will run an F-test to check if two continuous variables have the same means.

### Introduction

F test compares the means of two continuous variables. In other words it shows if their means were statistically different. We should be careful, while using the F test, because of the strict normality assumption, where strict means approximately normal ditribution is not enough to satisfy that.

### Normality assumption check (*Internet usage for educational purposes (hours per day)*)

The [*Shapiro-Wilk test*](http://en.wikipedia.org/wiki/Shapiro%E2%80%93Wilk_test), the [*Lilliefors test*](http://en.wikipedia.org/wiki/Lilliefors_test) and the [*Anderson-Darling test*](http://en.wikipedia.org/wiki/Anderson_Darling_test) help us to decide if the above-mentioned assumption can be accepted of the *Internet usage for educational purposes (hours per day)*.

|  |  |  |
| --- | --- | --- |
| Method | Statistic | p-value |
| Lilliefors (Kolmogorov-Smirnov) normality test | 0.2223 | 2.243e-92 |
| Anderson-Darling normality test | 42.04 | 3.31e-90 |
| Shapiro-Wilk normality test | 0.7985 | 6.366e-28 |

So, the conclusions we can draw with the help of test statistics:

* based on *Lilliefors test*, distribution of *Internet usage for educational purposes (hours per day)* is not normal
* *Anderson-Darling test* confirms violation of normality assumption
* according to *Shapiro-Wilk test*, the distribution of *Internet usage for educational purposes (hours per day)* is not normal

As you can see, the applied tests confirm departures from normality.

### Normality assumption check (*Age*)

The [*Shapiro-Wilk test*](http://en.wikipedia.org/wiki/Shapiro%E2%80%93Wilk_test), the [*Lilliefors test*](http://en.wikipedia.org/wiki/Lilliefors_test) and the [*Anderson-Darling test*](http://en.wikipedia.org/wiki/Anderson_Darling_test) help us to decide if the above-mentioned assumption can be accepted of the *Internet usage for educational purposes (hours per day)*.

|  |  |  |
| --- | --- | --- |
| Method | Statistic | p-value |
| Lilliefors (Kolmogorov-Smirnov) normality test | 0.17 | 6.193e-54 |
| Anderson-Darling normality test | 32.16 | 1.26e-71 |
| Shapiro-Wilk normality test | 0.8216 | 9.445e-27 |

So, the conclusions we can draw with the help of test statistics:

* based on *Lilliefors test*, distribution of *Age* is not normal
* *Anderson-Darling test* confirms violation of normality assumption
* according to *Shapiro-Wilk test*, the distribution of *Age* is not normal

As you can see, the applied tests confirm departures from normality.

*In this case it is advisable to run a more robust test, then the F-test.*

## Description

This template will run an F-test to check if two continuous variables have the same means.

### Introduction

F test compares the means of two continuous variables. In other words it shows if their means were statistically different. We should be careful, while using the F test, because of the strict normality assumption, where strict means approximately normal ditribution is not enough to satisfy that.

### The F-test

Here is the the result of the *F test* to compare the means of *Internet usage for educational purposes (hours per day)* and *Age*.

|  |  |  |
| --- | --- | --- |
| Method | Statistic | p-value |
| F test to compare two variances | 0.08618 | 3.772e-180 |

We can see from the table (in the p-value coloumn) that there is a significant difference between the means of *Internet usage for educational purposes (hours per day)* and *Age*.

## Description

This template will run an F-test to check if two continuous variables have the same means.

### Introduction

F test compares the means of two continuous variables. In other words it shows if their means were statistically different. We should be careful, while using the F test, because of the strict normality assumption, where strict means approximately normal ditribution is not enough to satisfy that.

### The F-test

Here is the the result of the *F test* to compare the means of *cyl* and *drat*.

|  |  |  |
| --- | --- | --- |
| Method | Statistic | p-value |
| F test to compare two variances | 11.16 | 1.461e-09 |

We can see from the table (in the p-value coloumn) that there is a significant difference between the means of *cyl* and *drat*.

This report was generated with [R](http://www.r-project.org/) (3.0.1) and [rapport](http://rapport-package.info/) (0.51) in *0.814* sec on x86\_64-unknown-linux-gnu platform.

