Master in Statistics and Information Management

Statistical Analysis

LU7 – ANOVA

Tutorial: One-way ANOVA with the Real Statistics Resource Pack

ANA CRISTINA COSTA

MARIA HENRIQUES

Tutorial, November 2021

Contents

List of figures

Introduction

Several Excel add-ins are available on the Internet, containing varied statistical tools, such as XLSTAT (https://www.xlstat.com) and the Real Statistics Resource Pack (www.real-statistics.com). It is suggested to install the latter as it is completely free.

This tutorial exemplifies the use of the Real Statistics Resource Pack to apply the tests associated with the ANOVA process[1] using the data from Example 1 of LU7. The Excel file **LU7_Examples.xlsx** shows the output results of the Real Statistics Resource Pack.

This tutorial was developed using Excel 365 version.

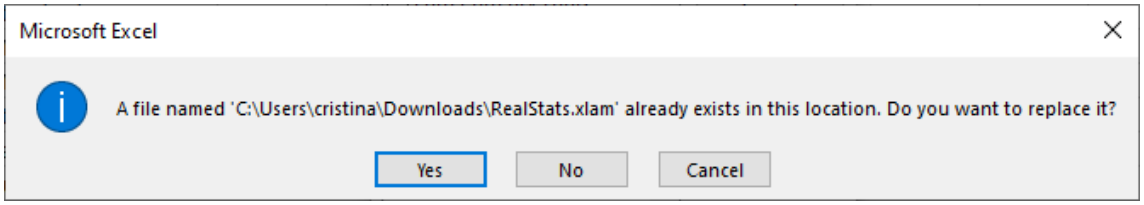
Installation of the Real Statistics Resource Pack

To install the Real Statistics Resource Pack (www.real-statistics.com), perform the following steps:

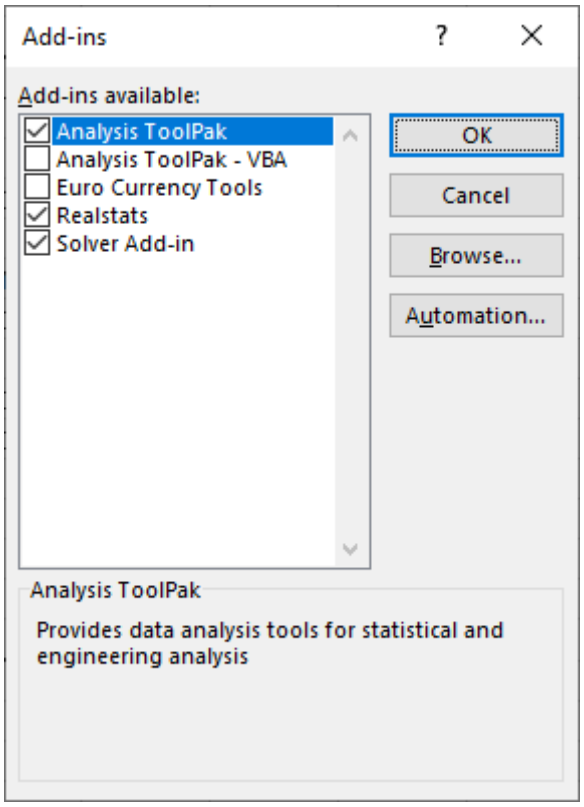
1. Download from the internet the Resource Pack file that is suitable for your operating system and Excel version (e.g., realstats.xlam) and save it in a folder on your computer[2]:
2. First, you need to activate the Solver add-in, as described below.
    - Click on the **File** menu.
    - Click on **Options**.
    - Click on **Add-ins** on the left panel; and
    - In the right panel, *View and Manage Microsoft Office Add-ins* box, click **Go..**.
    - Select the **Solver Add-in** option, and then click OK in all windows.
3. The Real Statistics Resource Pack add-in is activated, as described below.
    - Open the Excel workbook with the data you want to analyse.
    - Click on the **File** menu.
    - Click on **Options**.
    - Click on **Add-ins** on the left panel; and
    - In the right panel, *View and Manage Microsoft Office Add-ins* box, click **Go..**.
    - In the *Available Add-ins* window, click **Browse** to locate the file you downloaded from the internet (e.g., realstats.xlam), and then click OK.

[1] Costa, A. C. (2019). *One-way ANOVA process and nonparametric counterparts*. NOVA Information Management School, Universidade Nova de Lisboa, https://doi.org/10.13140/RG.2.2.25687.47520/1.
[2] The file can be saved in any folder in your computer, but later on you must remember which one it was. We recommend you to save the file in the Downloads folder.

- If a pop-up window, like the following one, appears, then click **Yes**:



- Finally, click OK in *the Add-ins* window:



The Real Statistics command will be available from the ADD-INS menu.

Important: You may need to repeat step 3 every time you initiate Excel.

Exploratory data analysis and normality testing

1. In the ADD-INS menu, open the Real Statistics > Data Analysis Tools, and then select the option **Descriptive Statistics and Normality**, and click Ok, as shown below (Figure 1Figure 2).

2. In the **Descriptive Statistics and Normality** window, place the cursor on the **Input Range** option and select the analysis data by dragging the mouse over the cells.

3. On the **options**, tick the boxes shown below (Figure 2).

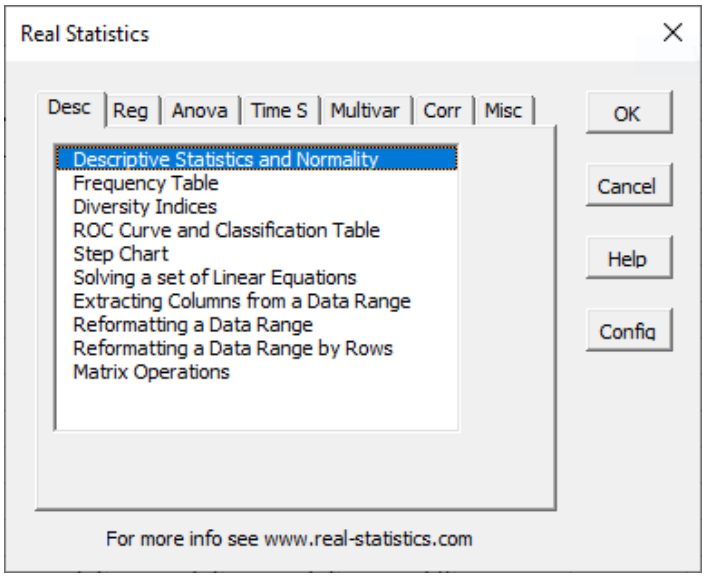


Figure 1. Request Descriptive Statistics and Normality tests

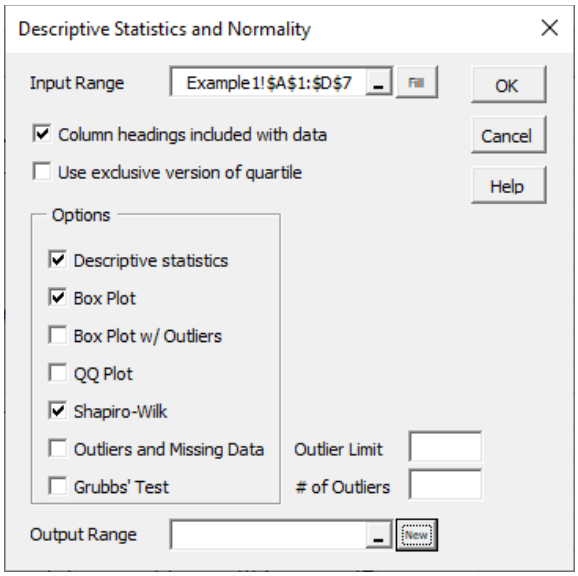


Figure 2. "Descriptive Statistics and Normality" window

ANOVA and associated tests

To perform the ANOVA and associated tests, follow the steps below. This add-in performs better if data is not staked in two columns (one for the response variable, and another one identifying each group), but rather with each sample in a different column (**Input format = Excel format**).

- In the ADD-INS menu, open the Real Statistics > Data Analysis Tools, select the **ANOVA** tab and then select the **One Factor Anova** option (Figure 3).

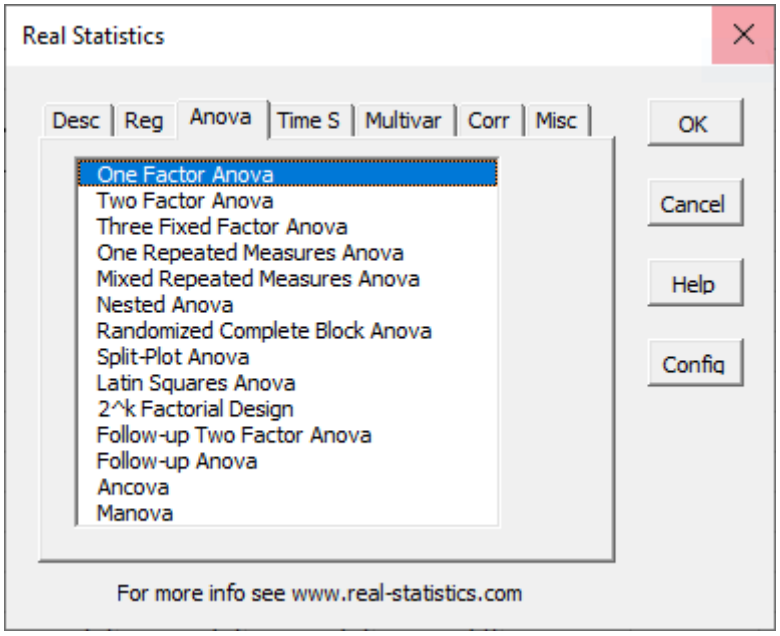


Figure 3. ANOVA tab in the "Real Statistics" window

- In the **ANOVA: Single Factor** window (Figure 4), select the data and appropriate ANOVA methods for the experimental design associated with the data:

  1. Place the cursor on the **Input Range** option and select the data by dragging the mouse over the cells. If you include the first row with the names of the variables, then the *Column headings included with data* option must be enabled.

     - "Excel format": when the data are organized so that the observations of each sample (group) are in different columns.

     - "Standard (stacked) format": when the data is organized so that all observations of the response variable are in the same column, and the first column identifies each group.

  2. Set the significance level (Alpha). Later, if you get an error, return to this menu, and use a comma to separate the decimal places.

In the **ANOVA: Single Factor** window, in addition to the ANOVA F-test, several statistical tests associated with ANOVA can be selected:

- Omnibus test options:
    - **Levene's test** to verify the assumption of homogeneity of variances (the output provides results of the test statistic based on both the means and medians of the samples of each group).
    - **Kruskal-Wallis test** (non-parametric).
    - **Welch Test** (for normal heteroskedastic populations) → **recommended**
    - **Brown-Forsythe test** (for normal heteroskedastic populations).

- Multiple comparison tests:
    - **Contrasts test** (used to perform contrast analysis[3]; a significance level correction must be used).
    - **Tukey's HSD test** (and **Tukey-Kramer's test**) → **recommended**
    - **Games-Howell test** (non-parametric; does not assume homoscedasticity).
    - **Dunnett test** (used to compare samples to a control group).
    - **Scheffée test** (used in complex comparisons; it is more conservative than the Tukey test but has the advantage of having some robustness regarding normality and homoscedasticity conditions).
    - **KW Contrasts test** (used to perform contrast analysis after the Kruskal-Wallis test; it is recommended to use a significance level correction)
    - **Nemenyi test** (non-parametric; identical to Tukey's HSD test; used after the Kruskal-Wallis test) → **recommended**
    - **Dunn test** (non-parametric; identical to the Nemenyi test but for samples of different sizes; used after the Kruskal-Wallis test)
    - **Dunnett KW test** (identical to the Dunnett test, but non-parametric; used to compare samples with a control group; used after the Kruskal-Wallis test)

- The **Alpha correction for contrasts** allows you to select the method you want to use to correct, or not (No Correction), the significance level of the multiple comparison tests. In the set of available tests, only Contrast tests require the use of Dunn-Sydak or Bonferroni correction.

[3] The contrast analysis is based on hypothesis tests on linear combinations of the mean values of the groups where the sum of the weights (constants) associated with the mean values is equal to zero. For more information, please consider the documentation on the Real Statistics Resource Pack supplement (http://www.real-statistics.com/one-way-analysis-of-variance-anova/unplanned-comparisons/; accessed on the 13th of November, 2021).

- In the **Output Range** field, select a cell to define the upper left cell of the results table, or click on the "Ne" button to send the results to a new sheet of the Excel file being used.

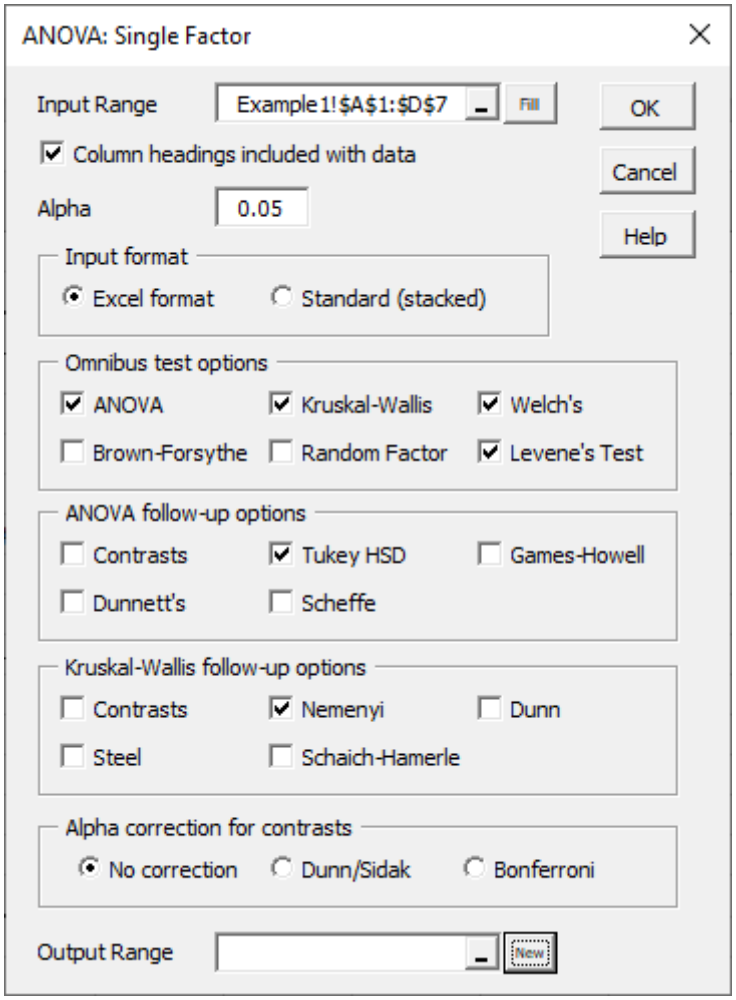


Figure 4. Statistical tests available in the "ANOVA: Single Factor" window

Nemenyi test

In the table of results of the Nemenyi test, a shaded column in gray will appear, called **c** (contrast), where the user must specify the weight associated with each mean value of the desired hypothesis test[4]. Considering the data in the example, let us suppose that we intend to test H_0: $\mu_B - \mu_C = 0$. In this case, the user must enter the value 1 in the cell corresponding to group B, and the value –1 in the cell corresponding to group C (Figure 5).

[4] For more information, please consider the documentation on the Real Statistics Resource Pack supplement (http://www.real-statistics.com/one-way-analysis-of-variance-anova/unplanned-comparisons/; accessed on the 13th of November, 2021).

NEMENYI TEST					
Groups	*c*	*R sum*	*n*		
A		33.5	5		
B	1	85	6		
C	-1	52.5	6		
D		82	5		
	0	32.5	5.454545		
Q TEST				Alpha	0.05
std err	*k*	*q-stat*	*q-crit*	*x-crit*	*p-value*
15.04314	4	2.160453	3.633	54.65172	0.421687

Figure 5. Results of the Nemenyi test comparing the mean of groups B and C.