

# EuroOffice Modeller

The functionalities of the Modeller extension can be grouped into three categories:

- I. Statistical methods
- II. Optimization
- III. Time series analysis

In the following all these functionalities shall be demonstrated.

## I. Statistical methods

### Descriptive statistics

In the following we shall demonstrate the workings of the Descriptive statistics module.

Amongst the variety of possible statistics we are able to calculate the following quantities:

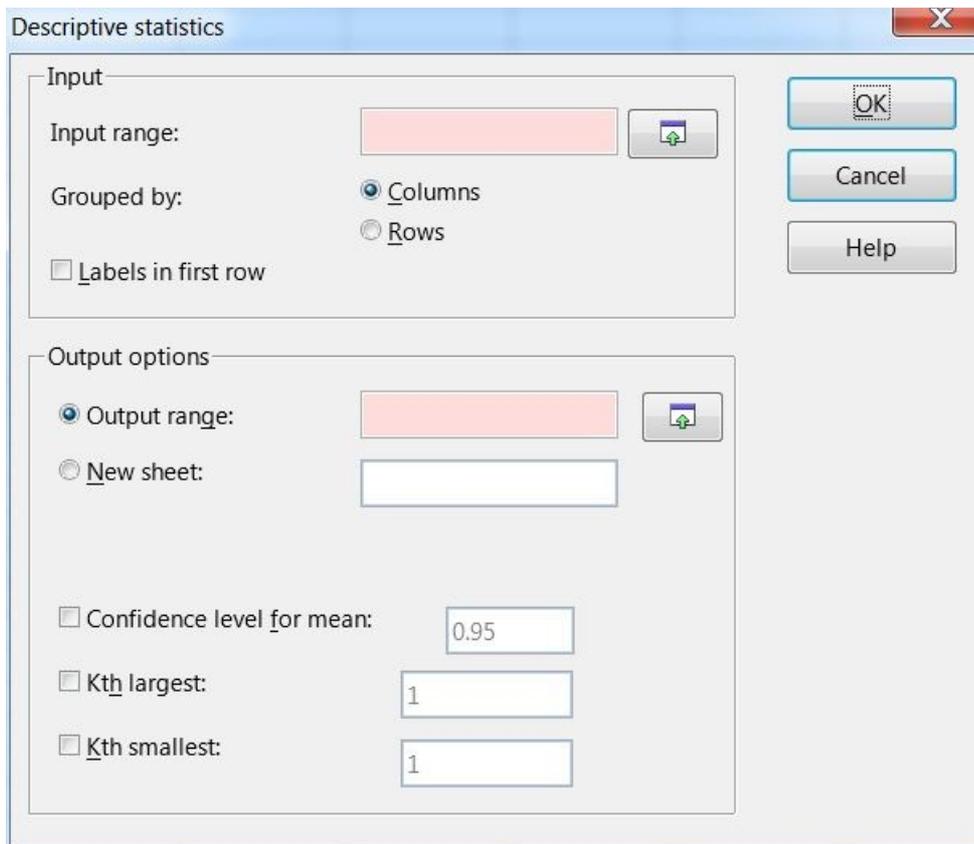
- Mean
- Standard error of the mean
- Median
- Mode
- Standard deviation
- Sample variance
- Skewness
- Range
- Minimum
- Maximum
- Sum
- Count
- Kurtosis

### Preparations

Let us open a spreadsheet and write in the dataset in question. The data can be in either rows or columns.

### Descriptive statistics

1. Let us click on the „**Modeller**” main menu and select the „**Descriptive statistics / Descriptive stats**” submenu, after which the dialog of the Descriptive Statistics module appears.



2. Click on the button next to the “Input range” textbox, after which we can select the data cell range. If there are labels we should select them as well and they will appear in the end result.
3. If the data are not in columns but rows then check the “Rows” option at the “Grouped by” buttons.
4. If the first row (or column in case the data is in columns) contains labels check the “Labels in first row” checkbox.
5. The “Output range” box contains the address of the output table (specifically its top left corner). We may select this by clicking on the button next to this field and select the required cell. Be careful of this selection since the output table will override every field so make sure there is enough space left under and to the right of this cell (the width of the table is seven fields while its height depends on the number of outputs chosen).
6. If we want to write the data to a new sheet then check the “New sheet” checkbox and enter the name of the new sheet. Be careful not to chose the name of an existing sheet!
7. The default value of “Confidence for mean” is set as 0.95. If we want to modify this check the checkbox near it and write the desired value in the field.
8. If we want to set the K-th largest value check the checkbox next to the “Kth largest” checkbox and enter the desired value in the field.
9. If we want to set the K-th smallest value check the checkbox next to the “Kth smallest” checkbox and enter the desired value in the field.
10. If we have finished with everything click the “OK” button, after which the result will appear shortly.

## Histogram

In the following we shall demonstrate the workings of the Histogram module of the EuroOffice Modeller extension. This calculates the frequency of the data in a specific bin and also displays this in a diagram.

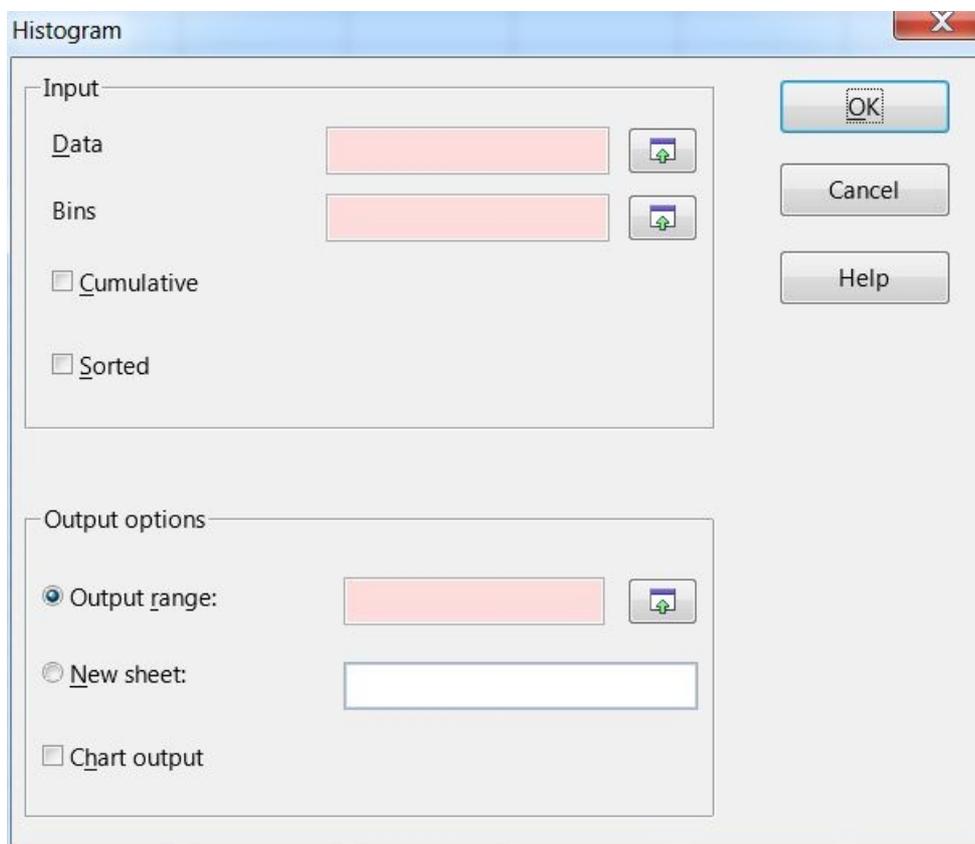
### Preparation

Write the data in a spreadsheet in EuroOffice calc. The data can be in either rows or columns. Two types of data are required:

- the data that need to be analyzed
- the boundaries of the bins

### Histogram

1. In the „Modeller” menu click on the „Descriptive statistics / Histogram” submenu, after which the dialog of the Histogram module appears.



2. Click on the button next to the “Data” field and select the data range with the help of the popup dialog box.
3. Click on the button next to the “Bins” field and select the bins range with the help of the popup dialog box. It is required that the numbers that appear here should be monotone decreasing otherwise the program will warn us of this.
4. If we would like to see the cumulative in the output check the “Cumulative” checkbox. The result will be displayed in percentages.
5. If we would like the program to sort the result check the “Sorted” checkbox. With this the result will also obtain the ordered results with their appropriate bin borders.

6. The “Output range” box contains the address of the output table (specifically its top left corner). We may select this by clicking on the button next to this field and select the required cell. Be careful of this selection since the output table will override every field so make sure there is enough space left under and to the right of this cell (the width of the table is seven fields while its height depends on the number of outputs chosen).
7. If we want to write the data to a new sheet then check the “New sheet” checkbox and enter the name of the new sheet. Be careful not to chose the name of an existing sheet!
8. If we would like to see a digram output check the “Chart output” checkbox.
9. If we have finished with everything click the “OK” button, after which the result will appear shortly.

## **Mean Tests**

Mean tests analyze the mean of scale variables in an attempt to provide significant differences between various groups. These are statistically correct only in the case where the distribution of the data are normal or they are almost normal. If the user is not certain whether the data are normally distributed it is worth checking this with either the Shapiro-test or using the kurtosis/skewness indicator of the descriptive statistics; from these one may find out whether it is appropriate to use mean tests on the data.

## **One-Sample Z-test**

In the following we shall demonstrate the workings of the One-Sample Z-test module of the EuroOffice Modeller extension.

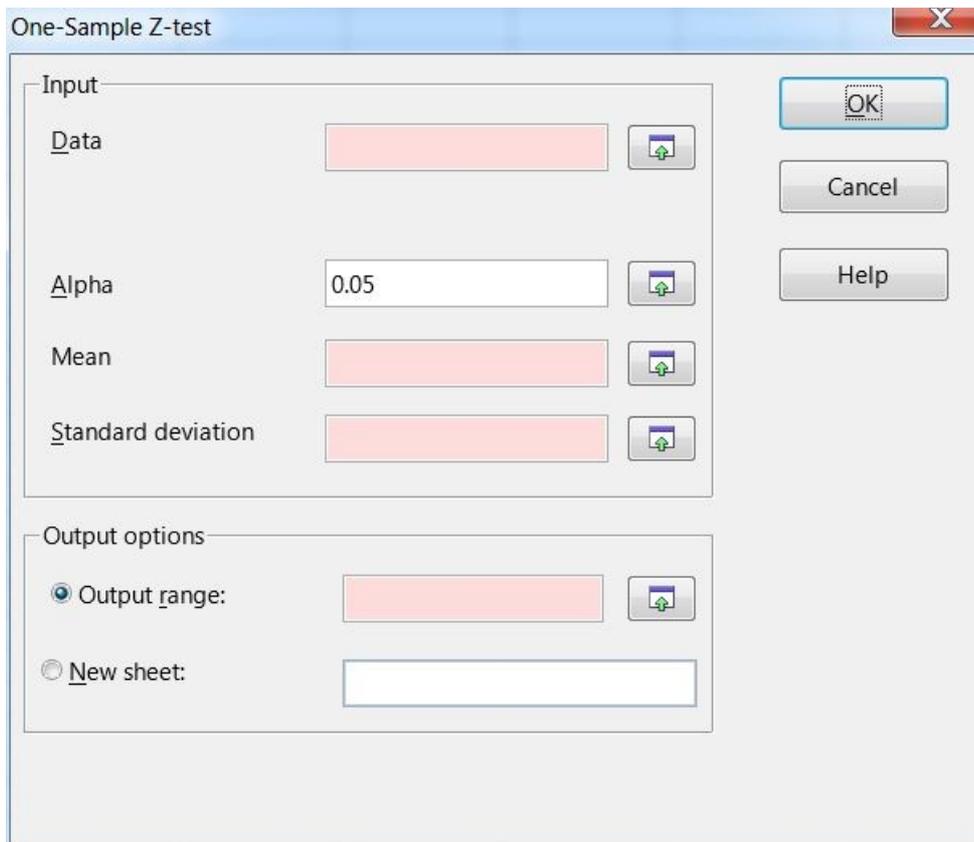
We compare the means of a normally distributed sample to an assumed mean with known standard deviation (this procedure is also known in the literature as u-test).

## **Preparation**

Write the data in a spreadsheet in EuroOffice calc. The data can be in either rows or columns.

## **One-Sample Z-test**

1. In the „**Modeller**” menu click on the „**Mean Tests / One-Sample Z-test**” submenu, after which the dialog of the One-Sample Z-test module appears.



2. We can provide the data in the “Data” text field. If we click on the button next to it we can select it with the appearing dialog.
3. The “Alpha” text field contains the significance value. Its default value is 0.05.
4. Write the value of the mean in the „Mean” text field.
5. Write the value of the standard deviation in the „Standard deviation” text field.
6. The “Output range” box contains the address of the output table (specifically its top left corner). We may select this by clicking on the button next to this field and select the required cell. Be careful of this selection since the output table will override every field so make sure there is enough space left under and to the right of this cell (the width of the table is seven fields while its height depends on the number of outputs chosen).
7. If we want to write the data to a new sheet then check the “New sheet” checkbox and enter the name of the new sheet. Be careful not to chose the name of an existing sheet!
8. If we have finished with everything click the “OK” button, after which the result will appear shortly.
9. The null hypothesis is that the sample is from a population which has the given mean. This can be rejected if  $p < 0.05$  (or other given significance value), that is the means differ significantly.

## T-test

In the following we shall demonstrate the workings of the T-test module of the EuroOffice Modeller extension.

With the help of the T-test we can analyze the means of normally distributed samples. We may work with

- independent samples

- paired samples. This can be set at the „**Test mode**” field. (One example for independent samples could be the results of the race between two different school classes; for the paired samples one example could be the results of a race of a school class at different times).

We may perform the calculations in two different types:

- equal variances
- non-equal variances

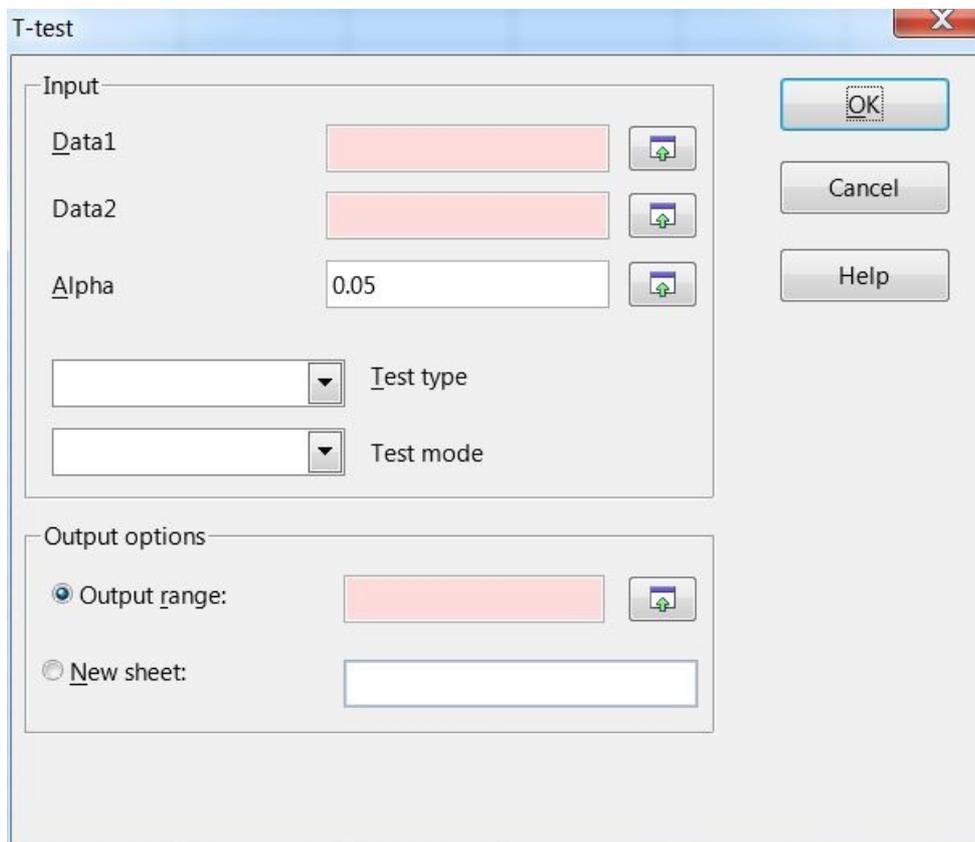
This can be given at the „**Test type**” field. If we are unsure about this we may perform an F-test before performing the T-test to check if the variances differ in a significant way. (The case where the samples are independent and the variances differ is also called Welch-test in the literature).

## Preparation

Write the data in a spreadsheet in EuroOffice calc. The data can be in either rows or columns.

## Independent unpaired T-test

1. In the „**Modeller**” menu click on the „**Mean Tests / T-test**” submenu, after which the dialog of the T-test module appears. At the **Test mode** field select the **Unpaired** case.



2. We can give the the data for the two samples in the „Data1” and „Data2” text fields. If we click on the button next to them we can select them with the appearing dialog.
3. The „Alpha” text field contains the value of the confidence level. Its default value is 0.05.
4. We can tell the program whether we have equal or unequal variances in the “Test type” text field.

It may happen that we do not have enough information. In this case one may perform an **F-test** before to make sure which method to chose.

- The “Output range” box contains the address of the output table (specifically its top left corner). We may select this by clicking on the button next to this field and select the required cell. Be careful of this selection since the output table will override every field so make sure there is enough space left under and to the right of this cell (the width of the table is seven fields while its height depends on the number of outputs chosen).
- If we want to write the data to a new sheet then check the “New sheet” checkbox and enter the name of the new sheet. Be careful not to chose the name of an existing sheet!
- If we have finished with everything click the “OK” button, after which the result will appear shortly.
- The null hypothesis is that the sample is from a population which has the given mean. This can be rejected if  $p < 0.05$  (or other given significance value), that is the means differ significantly.

### Paired T-test

- In the „Modeller” menu click on the „Mean Tests / T-test” submenu, after which the dialog of the T-test module appears. At the **Test mode** field select the **Paired** case.

- We can give the the data for the two samples in the „Data1” and „Data2” text fields. If we click on the button next to them we can select them with the appearing dialog.
- The „Alpha” text field contains the value of the confidence level. Its default value is 0.05.
- We can tell the program whether we have equal or unequal variances in the “Test type” text field. In the case of the paired test usually the equal variance is the preferred choice but if we are not sure we may decide it after performing an F-test.
- The “Output range” box contains the address of the output table (specifically its top left corner). We may select this by clicking on the button next to this field and select the required cell. Be careful of this selection since the output table will override every field so make sure there is

enough space left under and to the right of this cell (the width of the table is seven fields while its height depends on the number of outputs chosen).

6. If we want to write the data to a new sheet then check the “New sheet” checkbox and enter the name of the new sheet. Be careful not to chose the name of an existing sheet!
7. If we have finished with everything click the “OK” button, after which the result will appear shortly.
- 8 The null hypothesis is that the sample is from a population which has the given mean. This can be rejected if  $p < 0.05$  (or other given significance value), that is the means differ significantly.

### **ANOVA (Variance analyzis)**

In the following we shall demonstrate the workings of the ANOVA module of the EuroOffice Modeller extension. With the ANOVA method one can compare the means of a scale variable with the help of analyzing the variances, grouping them with respect to one or more nominal variables.

#### **Preparation**

Write the data in a spreadsheet in EuroOffice calc. The data can be in either rows or columns.

There are two types of tests which can be selected at the Methods / Type field:

- One-way
- Two-way

In the case of two-way ANOVA write the labels for the groups in the appropriate row (or column if the data are in columns).

#### **One-way ANOVA**

1. In the „Modeller” menu click on the „Mean Tests / ANOVA” submenu, after which the dialog of the ANOVA module appears.

ANOVA

Input

Input range:

Grouped by:  Columns  Rows

Labels in first row

Output options

Output range:

New sheet:

Methods

Type: One-way

Alpha: 0.05

Groups:

Sample size:

OK

Cancel

Help

2. Click on the button next to the “Input range” textbox, after which we can select the data cell range. If there are labels we should select them as well and they will appear in the end result.
3. If the data are not in columns but rows then check the “Rows” option at the “Grouped by” buttons.
4. If the first row (or column in case the data is in columns) contains labels check the “Labels in first row” checkbox.
5. The “Output range” box contains the address of the output table (specifically its top left corner). We may select this by clicking on the button next to this field and select the required cell. Be careful of this selection since the output table will override every field so make sure there is enough space left under and to the right of this cell (the width of the table is seven fields while its height depends on the number of outputs chosen).
6. If we want to write the data to a new sheet then check the “New sheet” checkbox and enter the name of the new sheet. Be careful not to chose the name of an existing sheet!
7. In the „Type” drop-down menu select the “One-way” option.
8. The „Alpha” text field contains the value of the confidence level. Its default value is 0.05. If we want to change it then we can either enter the value or click on the button next to the text field and select the cell which contains this value.
9. If we have finished with everything click the “OK” button, after which the result will appear shortly.

10. The result table consists of two main blocks. The first contains the general statistics of the groups while the second contains the main elements of the variance analysis. If  $p < 0.05$  then at least one of the means of the groups is significantly different from the others.

## Two-way ANOVA

1. In the „Modeller” menu click on the „Mean Tests / ANOVA” submenu, after which the dialog of the ANOVA module appears.

ANOVA

Input

Input range: [ ] [ ]

Grouped by:  Columns  Rows

Labels in first row

Output options

Output range: [ ] [ ]

New sheet: [ ]

Methods

Type: Two-way with repl. ▼

Alpha: 0.05 [ ] [ ]

Groups: [ ] [ ]

Sample size: [ ] [ ]

OK

Cancel

Help

2. In the „Type” drop-down menu select the “Two-way” option enabling extra dialog elements.
3. Click on the button next to the “Input range” textbox, after which we can select the data cell range. If there are labels we should select them as well and they will appear in the end result.
4. If the data are not in columns but rows then check the “Rows” option at the “Grouped by” buttons.
5. If the first row (or column in case the data is in columns) contains labels check the “Labels in first row” checkbox.
6. The “Output range” box contains the address of the output table (specifically its top left corner). We may select this by clicking on the button next to this field and select the required cell. Be careful of this selection since the output table will override every field so make sure there is enough space left under and to the right of this cell (the width of the table is seven fields while its height depends on the number of outputs chosen).
7. If we want to write the data to a new sheet then check the “New sheet” checkbox and enter the name of the new sheet. Be careful not to chose the name of an existing sheet!

8. The „Alpha” text field contains the value of the confidence level. Its default value is 0.05. If we want to change it then we can either enter the value or click on the button next to the text field and select the cell which contains this value.
9. Clicking on the button next to the “Groups” text field we can select the fields containing the groups.
10. Give the sample size in the “Sample size” text field (if this data appears on the spreadsheet then we may select it after clicking on the button next to this text field).
11. If we have finished with everything click the “OK” button, after which the result will appear shortly.
12. The result table consists of two main blocks. The first contains the general statistics of the groups while the second contains the main elements of the variance analysis. If  $p < 0.05$  then at least one of the means of the groups is significantly different from the others.

## Shapiro test

With this module we can check if the sample can be treated as one from a normally distributed data set.

1. From the „Modeller” menu click the „Normality Tests / Shapiro-test” submenu, after which the dialog of the Shapiro-test module appears.

2. Click on the button next to the “Input range” textbox, after which we can select the data cell range. If there are labels we should select them as well and they will appear in the end result.
3. The “Output range” box contains the address of the output table (specifically its top left corner). We may select this by clicking on the button next to this field and select the required cell. Be careful of this selection since the output table will override every field so make sure there is enough space left under and to the right of this cell.
4. If we have finished with everything click the “OK” button, after which the result will appear shortly.
5. Interpreting the results: the null hypothesis is that the sample is normally distributed. Depending on the p-value of the result we may accept or reject this hypothesis; if  $p < 0.05$  (the default significance level) then the difference is significant and the null hypothesis can be rejected and the sample is not normally distributed. If  $p > 0.05$  then we can keep the null hypothesis and the sample is normally distributed.

## Two-tailed F-test

In the following we shall demonstrate the workings of the Two-tailed F-test module of the EuroOffice Modeller extension.

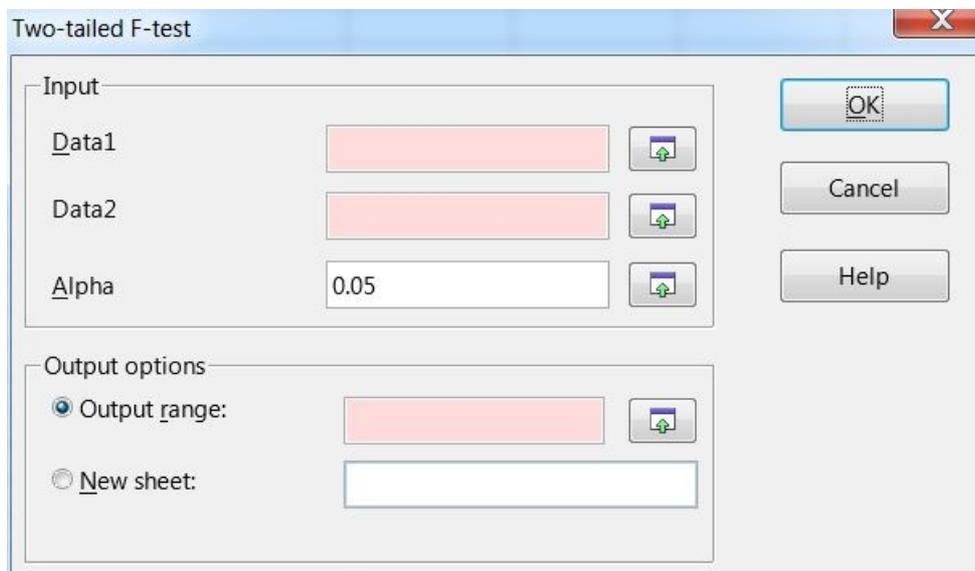
With this test we can compare the standard deviation of two samples if they are scale variables and normally distributed. The F-test also plays an important role in variance analysis and evaluating regression models. If the normality of the samples are unknown one should check this with either the Shapiro-test or with the Descriptive statistics module.

### Preparation

Write the data in a spreadsheet in EuroOffice calc. The data can be in either rows or columns.

### Two-tailed F-test

1. From the „Modeller” menu click the „Variance Tests / Two-tailed F-test” submenu, after which the dialog of the Two-tailed F-test module appears.



2. We can give the the data for the two samples in the „Data1” and „Data2” text fields. If we click on the button next to them we can select them with the appearing dialog.
2. The „Alpha” text field contains the value of the confidence level. Its default value is 0.05.
3. The “Output range” box contains the address of the output table (specifically its top left corner). We may select this by clicking on the button next to this field and select the required cell. Be careful of this selection since the output table will override every field so make sure there is enough space left under and to the right of this cell.
4. If we want to write the data to a new sheet then check the “New sheet” checkbox and enter the name of the new sheet. Be careful not to chose the name of an existing sheet!
5. If we have finished with everything click the “OK” button, after which the result will appear shortly.
6. The null hypothesis is that the two samples are from populations with the same variance. This can be rejected if  $p < 0.05$  (or other given significance level).

## Chi-squared test

In the following we shall demonstrate the workings of the Chi-squared test module of the EuroOffice Modeller extension.

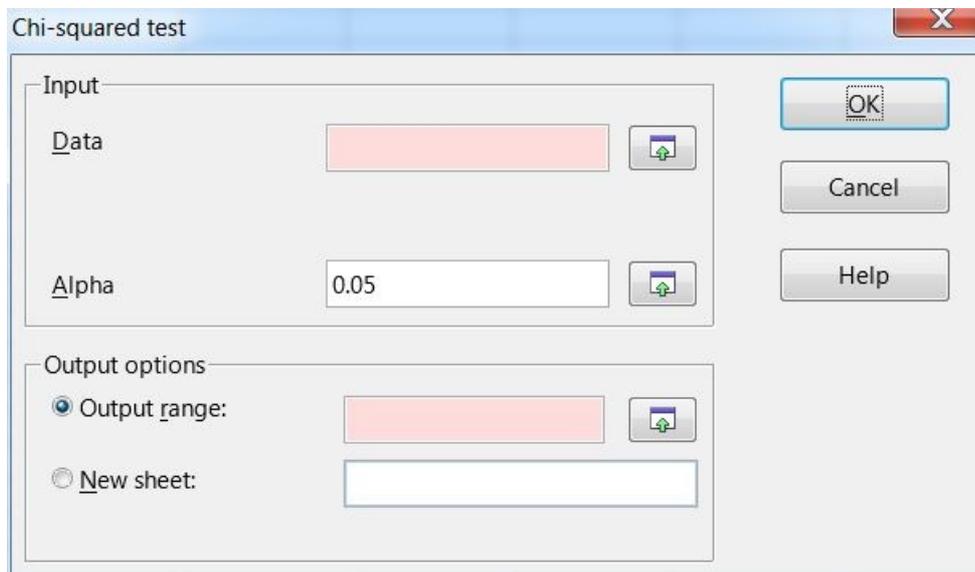
This module compares empirical data with some expected theoretical data. Its applications are quite widespread from checking independence to fitting tests and is also the starting point of cross table analysis.

### Preparation

Write the data in a spreadsheet in EuroOffice calc. The data can be in either rows or columns.

### Chi-squared test

1. From the „**Modeller**” menu click the „**Crosstable Analysis / Chi-squared test**” submenu, after which the dialog of the Chi-squared test module appears.



2. The observed data can be given in the „Data” textbox. By clicking on the button next to this field we can select the range with the help of the popup dialog. The expected values are calculated by the program assuming independency.
3. The „Alpha” text field contains the value of the confidence level. Its default value is 0.05.
4. The “Output range” box contains the address of the output table (specifically its top left corner). We may select this by clicking on the button next to this field and select the required cell. Be careful of this selection since the output table will override every field so make sure there is enough space left under and to the right of this cell.
5. If we want to write the data to a new sheet then check the “New sheet” checkbox and enter the name of the new sheet. Be careful not to chose the name of an existing sheet!
6. If we have finished with everything click the “OK” button, after which the result will appear shortly.
7. In case of checking independence the null hypothesis of the Chi-squared test is that the two quantities are independent. If  $p < 0.05$  (or the given significance level) then the null hypothesis can be rejected and the two quantities are dependent. If  $p > 0.05$  the null hypothesis can be accepted and the quantities are independent.

## Covariance

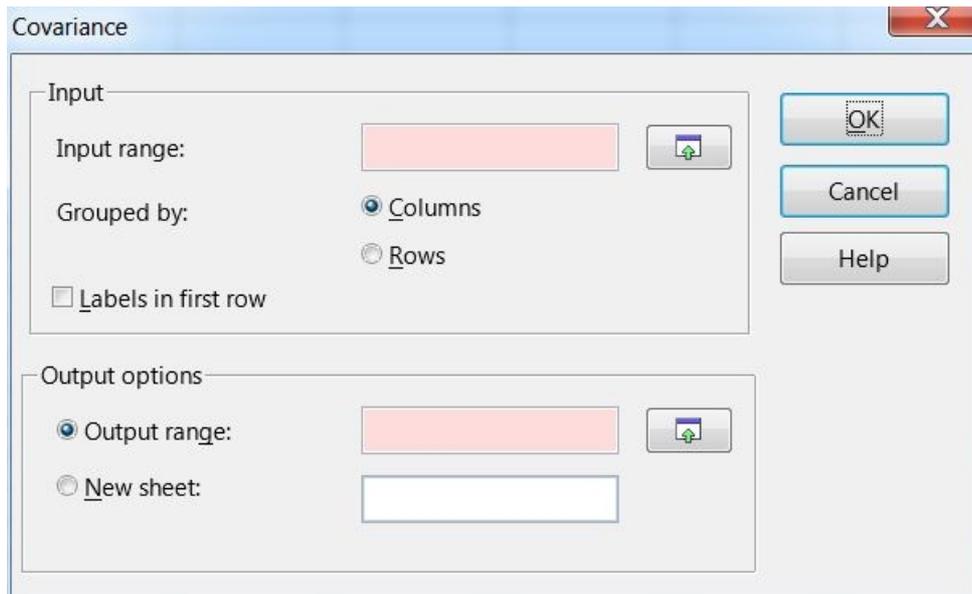
In the following we shall demonstrate the workings of the Covariance module of the EuroOffice Modeller extension. This module calculates the covariance matrix of scale variables.

### Preparation

Write the data in a spreadsheet in EuroOffice calc. The data can be in either rows or columns.

### Covariance

1. From the „**Modeller**” menu click the „**Covariance**” submenu, after which the dialog of the Covariance module appears.



2. Click on the button next to the “Input range” textbox, after which we can select the data cell range. If there are labels we should select them as well and they will appear in the end result.
3. If the data are not in columns but rows then check the “Rows” option at the “Grouped by” buttons.
4. If the first row (or column in case the data is in columns) contains labels check the “Labels in first row” checkbox.
6. The “Output range” box contains the address of the output table (specifically its top left corner). We may select this by clicking on the button next to this field and select the required cell. Be careful of this selection since the output table will override every field so make sure there is enough space left under and to the right of this cell.
7. If we want to write the data to a new sheet then check the “New sheet” checkbox and enter the name of the new sheet. Be careful not to chose the name of an existing sheet!
8. If we have finished with everything click the “OK” button, after which the result will appear shortly.

## Correlation

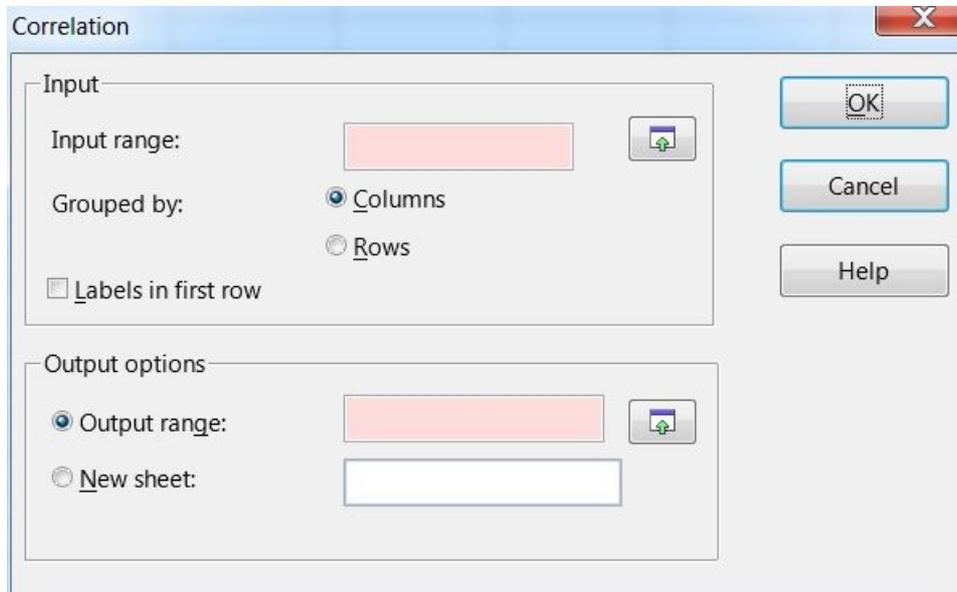
In the following we shall demonstrate the workings of the Correlation module of the EuroOffice Modeller extension. With this module one can make independency calculations by calculating the Pearson correlation  $\rho$  coefficient of scale variables.

## Preparation

Write the data in a spreadsheet in EuroOffice calc. The data can be in either rows or columns.

## Correlation

1. From the „**Modeller**” menu click the „**Correlation**” submenu, after which the dialog of the Covariance module appears.



2. Click on the button next to the “Input range” textbox, after which we can select the data cell range. If there are labels we should select them as well and they will appear in the end result.
3. If the data are not in columns but rows then check the “Rows” option at the “Grouped by” buttons.
4. If the first row (or column in case the data is in columns) contains labels check the “Labels in first row” checkbox.
5. The “Output range” box contains the address of the output table (specifically its top left corner). We may select this by clicking on the button next to this field and select the required cell. Be careful of this selection since the output table will override every field so make sure there is enough space left under and to the right of this cell.
6. If we want to write the data to a new sheet then check the “New sheet” checkbox and enter the name of the new sheet. Be careful not to chose the name of an existing sheet!
7. If we have finished with everything click the “OK” button, after which the result will appear shortly.
8. We will obtain the result as a symmetric table where the main diagonal consists of ones and the other elements are the Pearson-correlation of the appropriate rows and columns respectively.

## Regression

We can calculate the linear regression model of one dependent and one or more independent (explanatory) variables. This can be applied to scale variables.

### Regression analysis

1. From the „**Modeller**” menu click the „**Regression**” submenu, after which the dialog of the Regression module appears.

2. Click on the button next to the “Input range” text field. Then we can select the range containing the explanatory (independent) variables. If there are labels include them as well and they will appear in the end result.
3. The “Output range” box contains the address of the output table (specifically its top left corner). We may select this by clicking on the button next to this field and select the required cell. Be careful of this selection since the output table will override every field so make sure there is enough space left under and to the right of this cell.
5. In the “Y” text field we can give the Y (dependent) variable. We may also provide it by clicking the button next to the text field and select the appropriate range.
6. If we want a constant in the regression function (usually this is preferable) check the “Has constant” checkbox.
7. If we have finished with everything click the “OK” button, after which the result will appear shortly.
8. The table of the results consist of four main parts:
  - Regression analysis: the values and statistics of the coefficients
  - Model statistics: contains information and statistics of the power of the model
  - Residuals: statistics and normality analysis of the residuals
  - ANOVA-table: the characteristics of the model with respect to variance analysis; F-test; p-value

## Random

In the following we shall demonstrate the workings of the Random module of the EuroOffice Modeller extension. With the help of this module one can generate different a series of random numbers of various distributions. Implemented distributions:

### Beta-distribution

[https://en.wikipedia.org/wiki/Beta\\_distribution](https://en.wikipedia.org/wiki/Beta_distribution)

Alpha, Beta

The parameters of the distribution. Both have to be non-negative real numbers.

### Binomial distribution

[https://en.wikipedia.org/wiki/Binomial\\_distribution](https://en.wikipedia.org/wiki/Binomial_distribution)

n

The number of draws (integer).

p

The probability of the successful outcome, its value is between 0 and 1.

### Chi-squared distribution

[https://en.wikipedia.org/wiki/Chi-squared\\_distribution](https://en.wikipedia.org/wiki/Chi-squared_distribution)

df

The number of degrees of freedom.

### Custom distribution

In this case we can define our own distribution by providing the possible values and the probabilities associated to them.

Variables

The values which we want to choose from.

Probabilities

The probabilities for each value. The sum of them has to be one.

### Dirichlet-distribution

[http://en.wikipedia.org/wiki/Dirichlet\\_distribution](http://en.wikipedia.org/wiki/Dirichlet_distribution)

Alpha

The parameters (as a vector) of the distribution. All of them have to be greater than zero.

### Exponential distribution

[https://en.wikipedia.org/wiki/Exponential\\_distribution](https://en.wikipedia.org/wiki/Exponential_distribution)

Lambda

The rate/inverse scale parameter. Has to be greater than zero.

### F-distribution

<https://en.wikipedia.org/wiki/F-distribution>

df (Nominator )

Number of degrees of freedom in the nominator (can be any positive real number).

df (Denominator)

Number of degrees of freedom in the denominator (can be any positive real number).

### **Gamma-distribution**

[https://en.wikipedia.org/wiki/Gamma\\_distribution](https://en.wikipedia.org/wiki/Gamma_distribution)

Shape

The shape parameter of the distribution (positive, real).

Scale

The scale parameter of the distribution (positive, real).

### **Geometric distribution**

[https://en.wikipedia.org/wiki/Geometric\\_distribution](https://en.wikipedia.org/wiki/Geometric_distribution)

p

The probability of the successful outcome, its value is between 0 and 1.

### **Gumbel-distribution**

[https://en.wikipedia.org/wiki/Gumbel\\_distribution](https://en.wikipedia.org/wiki/Gumbel_distribution)

Location

The position parameter

Scale

The scale parameter, has to be a positive integer.

### **Hypergeometric distribution**

[https://en.wikipedia.org/wiki/Hypergeometric\\_distribution](https://en.wikipedia.org/wiki/Hypergeometric_distribution)

n

Number of successful draws.

m

Number of unsuccessful draws.

N

Number of drawn samples, its value has to be smaller than n+m.

### **Laplace-distribution**

[https://en.wikipedia.org/wiki/Laplace\\_distribution](https://en.wikipedia.org/wiki/Laplace_distribution)

Location

The position parameter

Scale

The scale parameter, has to be a positive integer.

### **Logistic distribution**

[http://en.wikipedia.org/wiki/Logistic\\_distribution](http://en.wikipedia.org/wiki/Logistic_distribution)

Location

The position parameter

Scale

The scale parameter, has to be a positive integer.

### **Log-normal distribution**

[https://en.wikipedia.org/wiki/Log-normal\\_distribution](https://en.wikipedia.org/wiki/Log-normal_distribution)

Mean

The expectation value of the distribution

Sigma

The standard deviation of the distribution (its value is a positive real number)

### **Logarithmic series distribution (also known as the Logarithmic or log-series distribution)**

[https://en.wikipedia.org/wiki/Logarithmic\\_distribution](https://en.wikipedia.org/wiki/Logarithmic_distribution)

p

The probability of the successful outcome, its value is between 0 and 1.

### **Multinomial distribution**

[https://en.wikipedia.org/wiki/Multinomial\\_distribution](https://en.wikipedia.org/wiki/Multinomial_distribution)

Experiments

The number of tries.

Probabilities

A vector containing the probabilities. All elements are summed to one.

### **Multivariate normal distribution**

[https://en.wikipedia.org/wiki/Multivariate\\_normal\\_distribution](https://en.wikipedia.org/wiki/Multivariate_normal_distribution)

Mean

The expectation value of the distribution

Covariance

The covariance matrix of the distribution. If the expectation value is a vector with n elements then the covariance matrix should be an n times n positive definite matrix.

### **Negative binomial distribution**

[https://en.wikipedia.org/wiki/Negative\\_binomial\\_distribution](https://en.wikipedia.org/wiki/Negative_binomial_distribution)

n

The number of tries. Its value have to be integer.

p

The probability of the successful outcome, its value is between 0 and 1.

### **Noncentral chi-squared distribution**

[https://en.wikipedia.org/wiki/Noncentral\\_chi-squared\\_distribution](https://en.wikipedia.org/wiki/Noncentral_chi-squared_distribution)

df

Number of degrees of freedom, a positive integer.

Non-centrality

The parameter describing non-centrality. Its value is greater than zero.

### **Noncentral F-distribution**

[https://en.wikipedia.org/wiki/Noncentral\\_F-distribution](https://en.wikipedia.org/wiki/Noncentral_F-distribution)

df (Nominator )

Number of degrees of freedom in the nominator (can be any positive real number).

df (Denominator)

Number of degrees of freedom in the denominator (can be any positive real number).

Non-centrality

The parameter describing non-centrality. Its value is greater than zero.

### **Normal distribution**

[https://en.wikipedia.org/wiki/Normal\\_distribution](https://en.wikipedia.org/wiki/Normal_distribution)

Mean

The expectation value of the distribution

Sigma

The standard deviation of the distribution. Its value is greater than zero.

### **Pareto-distribution**

Sampling from Pareto II or Lamax distribution. [https://en.wikipedia.org/wiki/Pareto\\_distribution](https://en.wikipedia.org/wiki/Pareto_distribution)

Shape

The shape parameter of the distribution. Its value is greater than zero.

### **Poisson-distribution**

[https://en.wikipedia.org/wiki/Poisson\\_distribution](https://en.wikipedia.org/wiki/Poisson_distribution)

Lambda

The lambda parameter of the distribution. Its value is greater than zero.

### **Power-distribution**

It's also called power law distribution. [https://en.wikipedia.org/wiki/Power\\_law#Power-law\\_probability\\_distributions](https://en.wikipedia.org/wiki/Power_law#Power-law_probability_distributions)

a

The parameter of the distribution. Its value is greater than zero.

### **Discrete uniform distribution**

[https://en.wikipedia.org/wiki/Discrete\\_uniform\\_distribution](https://en.wikipedia.org/wiki/Discrete_uniform_distribution)

Low

The lower bound of the interval.

High

The upper bound of the interval.

### **Rayleigh-distribution**

[https://en.wikipedia.org/wiki/Rayleigh\\_distribution](https://en.wikipedia.org/wiki/Rayleigh_distribution)

Scale

The scale parameter of the distribution. Its value is greater than zero.

### **Standard Cauchy-distribution**

[https://en.wikipedia.org/wiki/Cauchy\\_distribution](https://en.wikipedia.org/wiki/Cauchy_distribution).

The standard Cauchy-distribution is obtained when the expectation value is one and the scale parameter is zero.

### **Standard T-distribution**

[http://en.wikipedia.org/wiki/Student's\\_t-distribution](http://en.wikipedia.org/wiki/Student's_t-distribution)

df

Number of degrees of freedom, its value is greater than zero.

### **Triangular distribution**

[http://en.wikipedia.org/wiki/Triangular\\_distribution](http://en.wikipedia.org/wiki/Triangular_distribution)

Low

Lower bound.

High

Upper bound, its value is larger than the lower bound.

Mode

The peak of the distribution, its value has to be between the lower and the upper bounds.

### **Uniform distribution**

[https://en.wikipedia.org/wiki/Uniform\\_distribution\\_\(continuous\)](https://en.wikipedia.org/wiki/Uniform_distribution_(continuous))

Low

Lower bound of the interval.

High

Upper bound of the interval.

### **von Mises distribution**

[http://en.wikipedia.org/wiki/Von\\_Mises\\_distribution](http://en.wikipedia.org/wiki/Von_Mises_distribution)

Mu

The position parameter of the distribution.

Kappa

The dispersion parameter of the distribution. Its value is a non-negative real number.

### **Wald-distribution**

Also known as inverse Gaussian distribution.

[http://en.wikipedia.org/wiki/Inverse\\_Gaussian\\_distribution](http://en.wikipedia.org/wiki/Inverse_Gaussian_distribution)

Mean

The expectation value of the distribution which has to be positive.

Scale

The type of the distribution, its value is positive.

### **Weibull-distribution**

The general Weibull-distribution: [https://en.wikipedia.org/wiki/Weibull\\_distribution](https://en.wikipedia.org/wiki/Weibull_distribution) The distribution implemented is the case where the scale parameter is 1.

Shape

The shape parameter of the distribution.

### **Zipf-distribution**

<http://mathworld.wolfram.com/ZipfDistribution.html>

a

The parameter of the distribution.

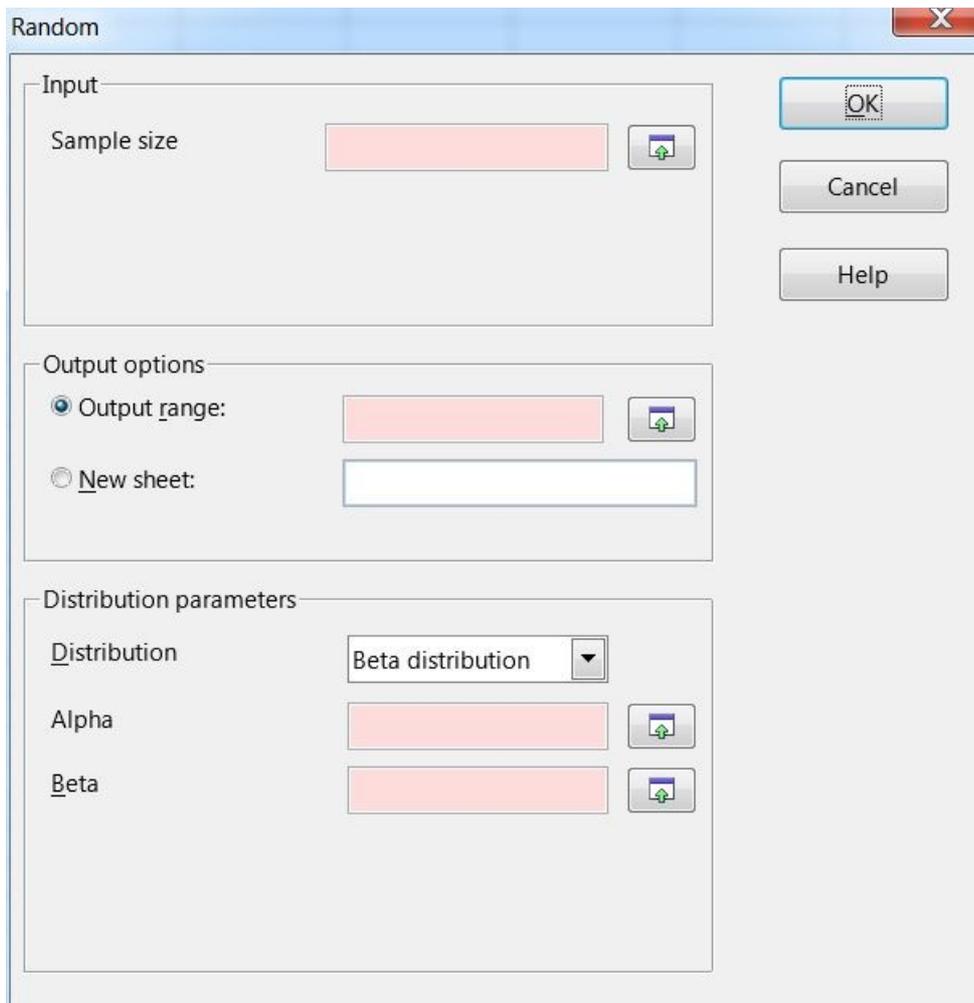
### **Preparation**

Let us open a new spreadsheet. If we wish the parameters of the distribution can be in the cells of the spreadsheet.

### **Modeller / Random**

#### **Random**

1. From the „**Modeller**” menu click the „**Random**” submenu, after which the dialog of the Random module appears.



2. We can give the sample size in the “Sample size” textbox. We may also give it by clicking on the button next to this text field and choose it from the spreadsheet with the help of the popup dialog.
3. The “Output range” box contains the address of the output table (specifically its top left corner). We may select this by clicking on the button next to this field and select the required cell. Be careful of this selection since the output table will override every field so make sure there is enough space left under and to the right of this cell.
4. If we want to write the data to a new sheet then check the “New sheet” checkbox and enter the name of the new sheet. Be careful not to chose the name of an existing sheet!
5. We can choose the type of the distribution in the “Distribution” dropdown menu. After clicking on it the parameters of the appropriate distribution will appear which have to be filled according to the requirements of that given distribution (for example in the case of the Binomial-distribution  $n$  is integer and  $p$  is between 0 and 1).
6. If we have finished with everything click the “OK” button, after which the result will appear shortly.

## Exponential smoothing

In the following we shall demonstrate the workings of the Random module of the EuroOffice Modeller extension. With the help of the module we may perform an exponential smoothing on a given dataset. The defining equations of the procedure are:

$$s_0 = x_0$$

$$s_t = \alpha x_{t-1} + (1 - \alpha)s_{t-1}, t > 0$$

## Preparation

Write the data in a spreadsheet in EuroOffice calc. The data can be in either rows or columns.

## Exponential smoothing

1. From the „**Modeller**” menu click the „**Exponential smoothing**” submenu, after which the dialog of the Exponential smoothing module appears.

Exponential Smoothing

Input

Input range:

Damping factor:

Grouped by:  Columns  Rows

Labels in first row

Output options

Output range:

New sheet:

Chart output

OK Cancel Help

2. Click on the button next to the “Input range” textbox, after which we can select the data cell range. If there are labels we should select them as well and they will appear in the end result.
3. The damping factor can be given in the “Damping factor” text field. Its value has to be between 0 and 1.
4. If the data are not in columns but rows then check the “Rows” option at the “Grouped by” buttons.
5. If the first row (or column in case the data is in columns) contains labels check the “Labels in first row” checkbox.
6. The “Output range” box contains the address of the output table (specifically its top left corner). We may select this by clicking on the button next to this field and select the required cell. Be careful of this selection since the output table will override every field so make sure there is enough space left under and to the right of this cell.
7. If we want to write the data to a new sheet then check the “New sheet” checkbox and enter the name of the new sheet. Be careful not to chose the name of an existing sheet!

8. If we want a diagram output as well check the “Chart output” checkbox.
9. If we have finished with everything click the “OK” button, after which the result will appear shortly.

## **II. Optimization**

### **General description**

The solver is capable of numerically solving maximization or minimization problems that involve functions – also referred to as target functions - depending on many variables and the domain of the problem is given by a set of inequalities. It can help with problems occurring in economic and financial fields, and is a useful tool in calculating problems in higher education and science e.g. optimal calculation of the cost of transportation, calculating economically sound investments etc.

If the target function is a linear function of the variables then we can use linear, otherwise non-linear optimizers.

EuroOffice contains its own developed solver which can also be downloaded separately from the web-page of MultiRacio Ltd. [www.multiracio.com](http://www.multiracio.com)

The EuroOffice Solver is a revised solver which is capable of solving both linear and non-linear problems and is able to handle both linear and non-linear constraints.

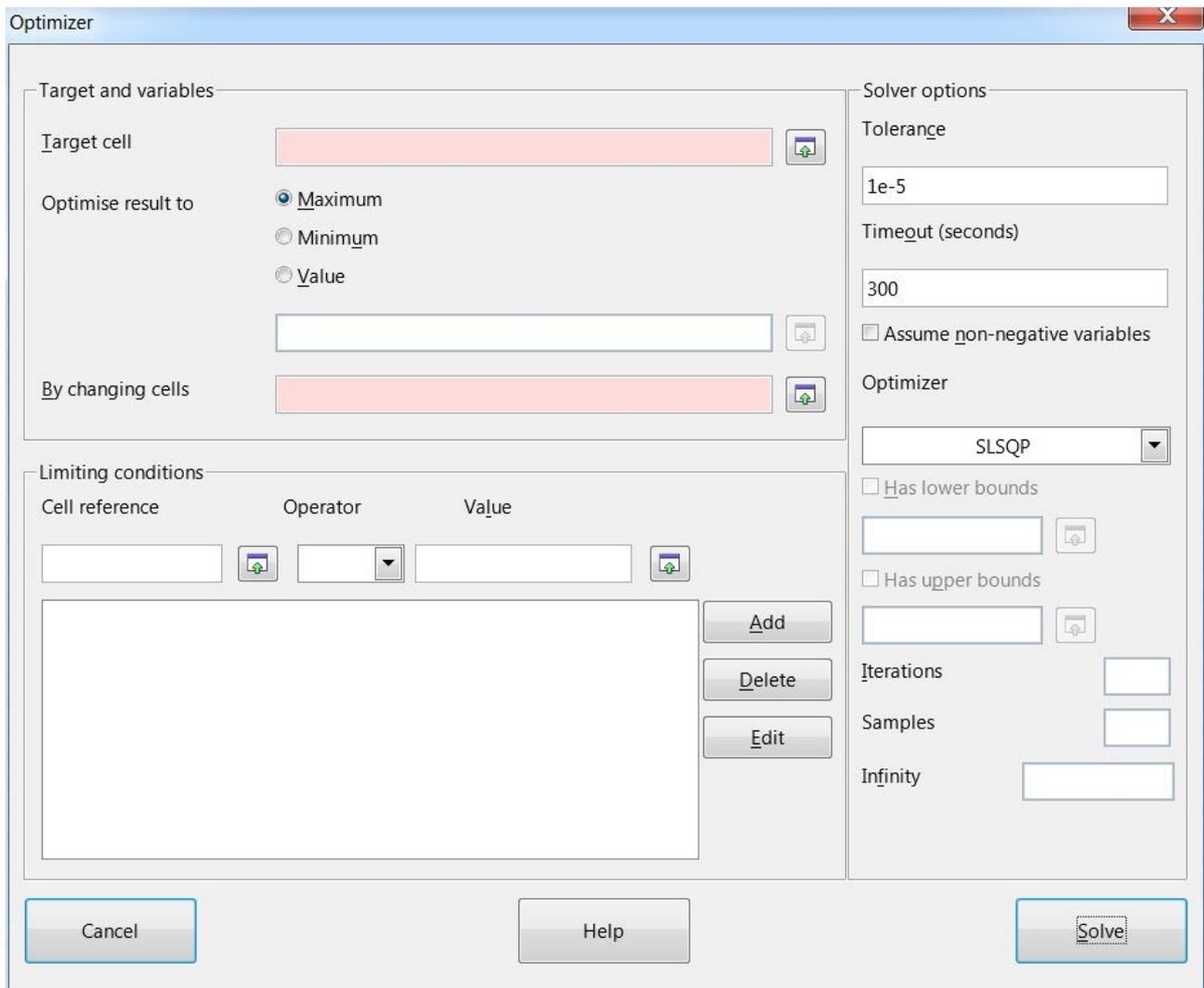
The EuroOffice Solver is an EuroOffice extension developed by Multiráció Ltd. (Budapest, Hungary).

### **EuroOffice Solver**

In the following we shall demonstrate the workings of the Solver module of the EuroOffice Modeller extension.

#### **Non-linear solver**

1. From the „**Modeller**” menu click the „**EuroOffice Solver**” submenu, after which the dialog of the EuroOffice Solver module appears.



The elements of the dialog are the following:

### Target cell

This contains the value of the target function which depends on the values of the “changing cells” (see later). We can give this value easily with the help of the button next to the text field.

### Optimise result to

Here we can decide if we want to minimize or maximize the target function or set it to a certain value.

### By changing cells

This contains the cells which are to be changed, that is those cells which the target cell is dependent of. If there are more changing cells then this can be a cell range.

### Limiting conditions

Here we can give the limiting conditions (constraints) of the problem if there are any. The left hand side of the limiting condition contains the cell or cell range of the function of the constraint. The right hand side contains either a value or a cell referring to a value. The possible choices of the Operator are the following:  $\leq$ ,  $=$ ,  $\geq$ , Integer, Binary. If we choose one of the latter two the right hand side has to be empty. The Integer condition means that the given function has to be integer while the Binary option means that the values of the given function can be either TRUE (1) or

FALSE (0).

## **Solver options**

We can change some of the solver settings on the right hand side of the dialog.

### **Tolerance**

This tolerance value is the criteria for the solver to stop., that is the calculation stops if the target function doesn't change at least the amount of this value. If it is too small then the obtained result might not be accurate so if we find the result to be strange then repeat the calculation with a different tolerance value.

### **Timeout (seconds)**

Some problems might be time consuming. If we don't want to wait for a long time then set this value appropriately.

### **Assume non-negative variables**

If all of the changing variables are positive then we can tell that to the program by checking this checkbox. This we we don't have to set this condition at the constraints.

### **Optimizer**

Here we may choose from the different types of optimization methods. We have the following possibilities:

**Linear Solver** In case we chose the linear solver, the target function and the constraints all have to be linear functions of the changing variable. The engine of the solver is the **lpsolve** package. In practice a lot of optimization problems are linear (e.g. assembly of products, transportation, time schedules etc. ) If we choose the linear solver the program checks whether the problem is actually linear and if not we will receive an error message. In case of linear problems we have the option to perform a **Sensitivity Analysis**.

**Nelder-Mead** With the help of the Nelder-Mead method we can solve non-linear problems which do not have constraints. This algorithm provides the next starting point by extrapolating the objective function: [http://en.wikipedia.org/wiki/Nelder-Mead\\_method](http://en.wikipedia.org/wiki/Nelder-Mead_method)

**SLSQP** With the help of the SLSQP method (Sequential Least Squares Programming) we can solve non-linear problems which have constraints. The method calculates the next starting point with the help of the objective function, the constraints and all their derivatives:: <http://www.pyopt.org/reference/optimizers.slsqp.html>. This is the **default** solver.

**NSGA2** The NSGA2 method is an evolution optimizer. It is useful in cases where either the target function or the constraints are not continuous. If we chose this optimizer we may set some extra parameters which can be found on the right hand side of the dialog (see below for details).

#### **Has lower bounds**

In case of the NSGA2 optimizer it is sometimes useful to restrict the bounds of the calculation. If we want to set the lower bounds manually, check the checkbox and chose a cell range. This range must be of the same shape as the changing cells and must contain a specific value for each variable. If it is left empty, all bounds are treated as if they wer the given infinity value (see below).

#### **Has upper bounds**

The same as the lower bound case, but this range is for the upper bounds for the variables.

## Iterations

The number of iterations. WARNING! If the problem has a lot of variable, then settings this value too large may result in a long calculation!

## Samples

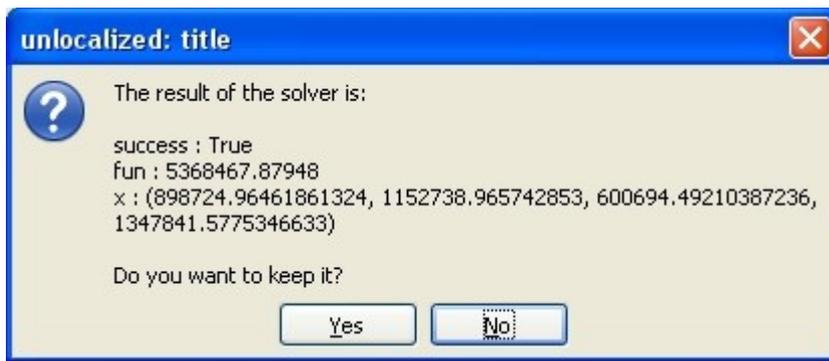
The number of samples in each iteration. WARNING! If the problem has a lot of variable, then settings this value too large may result in a long calculation!

## Infinity

The numerical value for the largest possible value of all changing cells. The larger it is, the more samples and iterations it takes the calculation to converge.

2. Set the value of the target cell in the “Target cell” text field. With the help of the button next to this field we can select it from the sheet as well.
3. Select the desired optimization type with the “Optimise result to” buttons. We may choose to minimize or maximize the target function or set it to a preferred value. In the latter case we can either give the specific value or a cell reference to it.
4. Set the changing cells in the “By changing cells” text field.
5. At the limiting conditions set the “Cell reference” to the cell reference containing the formula for the constraint.
6. Set the required operator for the constraint in the „Operator” dropdown menu. The possible choices of the Operator are the following:  $\leq$ ,  $=$ ,  $\geq$ , Integer, Binary. If we choose one of the latter two the right hand side has to be empty. The Integer condition means that the given function has to be integer while the Binary option means that the values of the given function can be either TRUE (1) or FALSE (0).
7. The “Value” field should contain the value appearing on the right hand side of the constraint.
8. At the right hand side of the dialog we can change some of the optimization settings.
9. „Assume non-negative variables” - If all of the changing variables are positive then we can tell that to the program by checking this checkbox. This we don't have to set this condition at the constraints.
10. „Tolerance” - This tolerance value is the criteria for the solver to stop., that is the calculation stops if the target function doesn't change at least the amount of this value. If it is too small then the obtained result might not be accurate so if we find the result to be strange then repeat the calculation with a different tolerance value.
11. „Timeout (seconds)” Some problems might be time consuming. If we don't want to wait for a long time then set this value appropriately.
12. From the „Optimizer” dropdown menu we may choose the type of optimizer we wish to use. If we don't have a linear problem we should consider using the default SLSQP method. If our problem is linear then we should use the linear solver. A short description of the possible methods have been provided earlier.
12. If we have finished with everything click the “Solve” button and wait for the calculation to finish. If the solver finds a solution satisfying our requirements it will ask if we want to accept it or not. If it doesn't find a suitable solution it will inform us about it as well.

After we have provided every information we can start the calculation by clicking on the **Solve** button. Depending on the model the calculation should finish within a few seconds and if the program has found a solution a popup message will appear, similar to the one below:



If we decide to keep the result click on the **Yes** button, otherwise click on the **No** button. Choosing the latter will result in the resetting the changing variables to their original value.

### III. Time series analysis

Within the **Time series analysis** module there are several methods which are available as a separate statistical function:

- Moving average
- Regression
- Ljung-Box test
- Shapiro-test

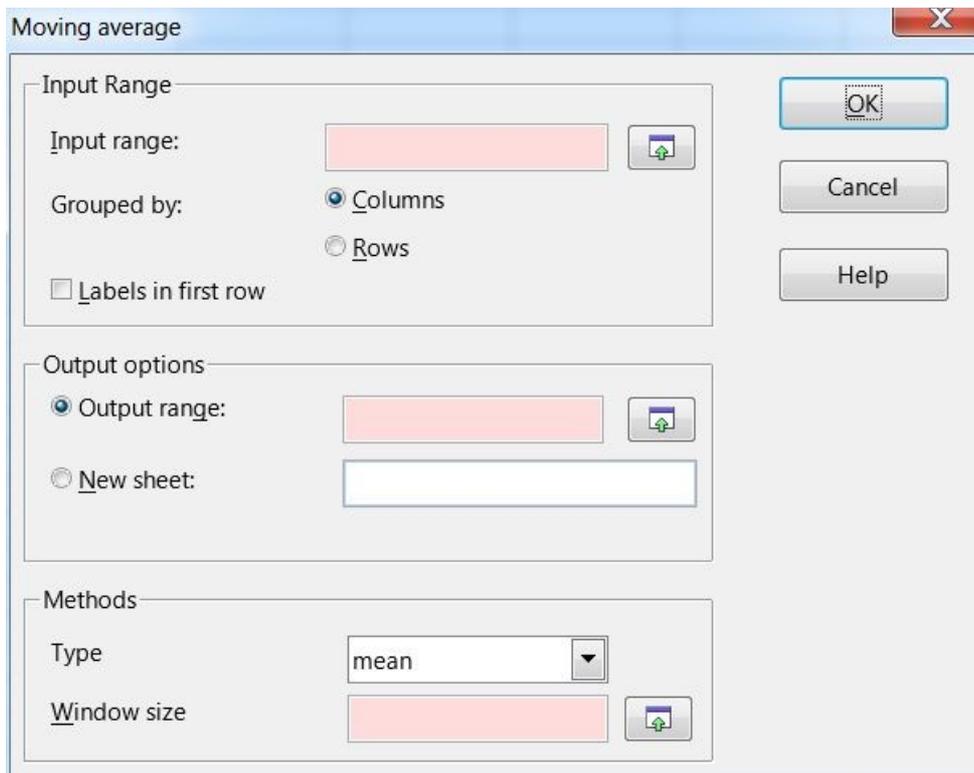
Choosing either of these methods will result in the appearance of the same dialog as choosing them from the other menus thus their functionalities remain the same and will not be repeated here.

In the following we shall demonstrate the workings of the **Time series analysis** module of the EuroOffice Modeller extension.

#### Moving average

This method is useful for smoothing data series and eliminating systematic fluctuations. It's applied mainly in case of time series analysis, it is used in many time series analysis module and it also plays a part in seasonal adjustments.

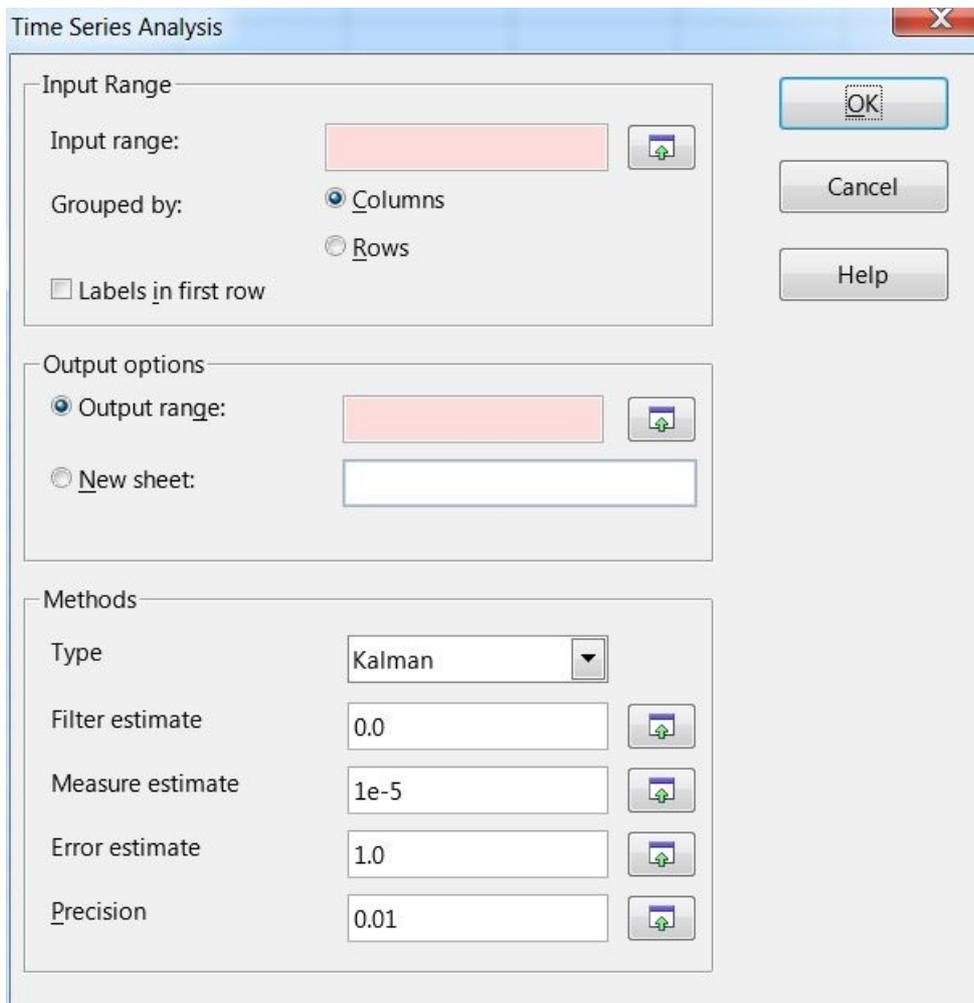
1. From the „**Modeller**” menu click the „**Timeseries Analysis / Moving average**” submenu, after which the dialog of the Moving average module appears.



2. In the „Type” dropdown menu set the “average” or the “sum” item. Choosing the former gives us the real moving average while choosing the sum will result in a moving sum for the data prior to the window size.
3. Click on the button next to the “Input range” textbox, after which we can select the data cell range. If there are labels we should select them as well and they will appear in the end result.
4. The “Output range” box contains the address of the output table (specifically its top left corner). We may select this by clicking on the button next to this field and select the required cell. Be careful of this selection since the output table will override every field so make sure there is enough space left under and to the right of this cell.
5. Set the period for which the program should calculate in the „Window size” text box.
6. If we have finished with everything click the “OK” button, after which the result will appear shortly.
7. The data column of the result will contain values Not Available (Na) while we reach a point which fits the window size. E.g. if the size of the window is 4 then the first three.

### **Kalman-filter**

1. From the „**Modeller**” menu click the „**Timeseries Analysis / TimeSeries**” submenu, after which the dialog of the Time series analysis module appears.
2. In the “Type” dropdown menu select “Kalman” after which extra dialog elements will appear.



3. Click on the button next to the “Input range” textbox, after which we can select the data cell range. If there are labels we should select them as well and they will appear in the end result.
4. The “Output range” box contains the address of the output table (specifically its top left corner). We may select this by clicking on the button next to this field and select the required cell. Be careful of this selection since the output table will override every field so make sure there is enough space left under and to the right of this cell.
5. At the „Filter estimate” text field we can set the estimate of the filter. We can also set it by clicking on the button next to this text field and choosing it from the spreadsheet.
6. At the „Measure estimate” text field we can set the estimate of the measurement. We can also set it by clicking on the button next to this text field and choosing it from the spreadsheet.
7. At the „Error estimate” text field we can set the estimate of the error of the measurement. We can also set it by clicking on the button next to this text field and choosing it from the spreadsheet.
8. At the „Precision” text field we can set the precision of the filter. We can also set it by clicking on the button next to this text field and choosing it from the spreadsheet.
9. If we have finished with everything click the “OK” button, after which the result will appear shortly.

### Ljung-Box test

1. From the „Modeller” menu click the „Timeseries Analysis / TimeSeries” submenu, after which the dialog of the Time series analysis module appears.

2. In the “Type” dropdown menu select “Ljung-Box” after which extra dialog elements will appear.

Time Series Analysis

Input Range

Input range:

Grouped by:  Columns  Rows

Labels in first row

Output options

Output range:

New sheet:

Methods

Type: Ljung-Box

Lag:

OK

Cancel

Help

3. Click on the button next to the “Input range” textbox, after which we can select the data cell range. If there are labels we should select them as well and they will appear in the end result.
4. The “Output range” box contains the address of the output table (specifically its top left corner). We may select this by clicking on the button next to this field and select the required cell. Be careful of this selection since the output table will override every field so make sure there is enough space left under and to the right of this cell.
5. Set the lag at the “Lag” text field. If this is zero then the program will perform the calculation for all lags.
6. If we have finished with everything click the “OK” button, after which the result will appear shortly.