

Program STATIST

<http://statist.wald.intevation.org/>

statist by Dirk Melcher

statist is a terminal-based statistics program with an interactive menu that makes it very easy to use. It can also be run in scripts and big datasets are handled reasonably well even on small machines. In spite of its low overhead statist can do quite a bunch of regression functions and tests. It can produce colorized output and uses gnuplot to create graphics.

Check the available [screenshots](#), [menu structure](#), [user manual](#), and a html view of statist [running](#) to see the possibilities.

Mailinglist: [Archive](#) | [Infos/subscribe](#)

Status of Translations (with links to menus)

Language	% translated	Last translator
German	100	Michael Gebhardt
Italian	93	Sandro Tosi
Portuguese	100	Jakson Aquino
Spanish	80	Carleos Artime

STATIST: Screenshots

```

Terminal
File Edit Settings Help

0 = Main menu
1 = Linear regression and correlation
2 = SPEARMAN rank-correlation-coefficient
3 = Multiple linear correlation
4 = Partial linear correlation (maximum: 5 variables)
5 = Polynomial regression
6 = Matrix of the linear correlation coefficients
7 = Matrix of SPEARMAN correlation coefficients
8 = Point-biserial (linear) correlation
9 = Cross-validation of multiple linear regression
10 = Randomization of multiple linear regression

Your choice: 3

Please enter number of columns to be correlated,
First selected column is taken as y-value: 3
Columns: a b c d e f g h i j k l m

Column for variable 1: a
Column for variable 2: b
Column for variable 3: c

y = Column a
x[1] = Column b
x[2] = Column c

Results multiple linear regression:
Regressed equation: y = B[0] + B[1]*x[1] + B[2]*x[2] + ... + B[n]*x[n]

B[0] = 36662,005276
B[1] = -2210,398571
B[2] = 1919,723672
                    
```

The Gnuplot window displays a 3D plot titled "STATIST: Multiple Linear Regression". The vertical axis is labeled 'a' and ranges from -150,000 to 250,000. The horizontal axes are labeled 'b' and 'c', both ranging from 0 to 120. A blue grid represents the regression plane, and red data points are scattered around it, showing a positive correlation between variables b and c, and a negative correlation between variable a and both b and c.

Statist - menu

- Main menu
 - 0 = Quit
 - 1 = Data management
 - 2 = Regressions and correlations
 - 3 = Tests
 - 4 = Miscellaneous
 - 5 = Data manipulation
 - 6 = Preferences

Statist - menu

- Preferences (volba 6)

- 0 = Main menu
- 1 = Save preferences
- 2 = Verbose [yes]
- 3 = Gnuplot graphics [yes]
- 4 = Beep at errors and warnings [yes]
- 5 = Histogram as text graphic instead of gnuplot-graphic [no]
- 6 = Special output changes from Bernhard [no]
- 7 = Use system command "ls" [no]
- 8 = Use value labels [no]
- 9 = Maximum number of rows before aborting table printing [200]
- 10 = Screen number of columns [95]
- 11 = Screen number of lines [34]

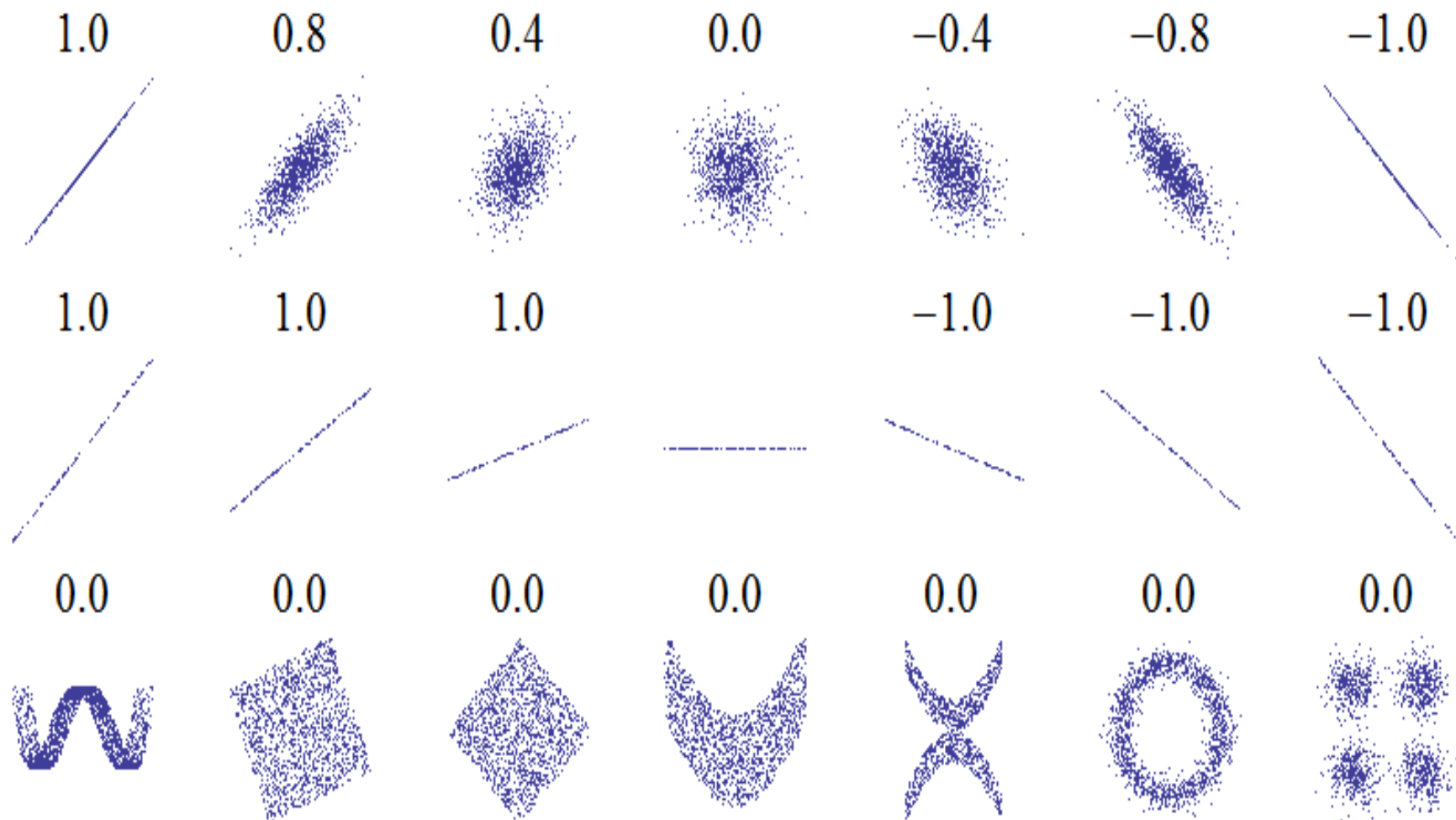
Statist - menu

- Data management (volba 1)
 - 0 = Main menu
 - 1 = List data of columns
 - 2 = Read another file
 - 3 = List names of columns
 - 4 = Rename column
 - 5 = Read column from terminal
 - 6 = Export columns as ASCII-data
 - 7 = Export data base as fixed width data file
 - 8 = File format options

Statist - menu

- Regrese a korelace (volba 2)
 - 0 = Main menu
 - 1 = Linear regression and correlation
 - 2 = SPEARMAN rank-correlation-coefficient
 - 3 = Multiple linear correlation
 - 4 = Partial linear correlation (maximum: 5 variables)
 - 5 = Polynomial regression
 - 6 = Matrix of the linear correlation coefficients
 - 7 = Matrix of SPEARMAN correlation coefficients
 - 8 = Point-biserial (linear) correlation
 - 9 = Cross-validation of multiple linear regression
 - 10 = Randomization of multiple linear regression

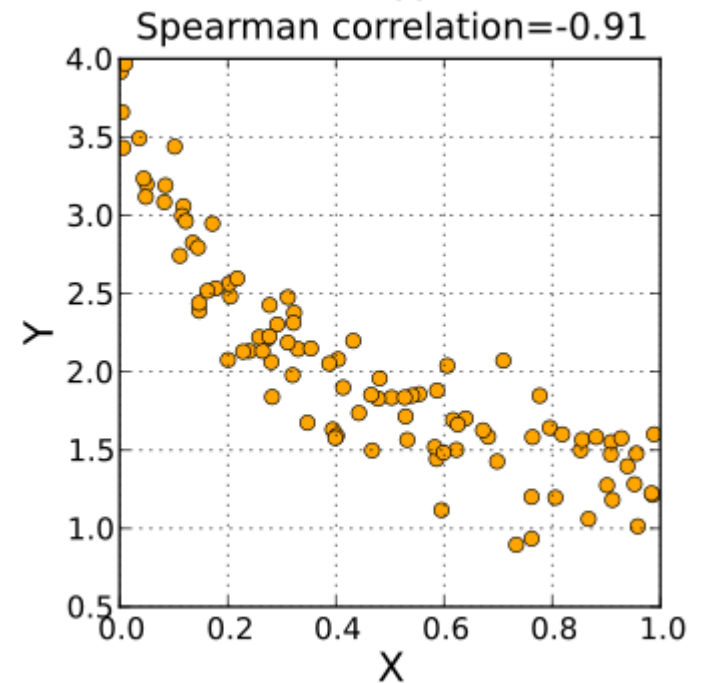
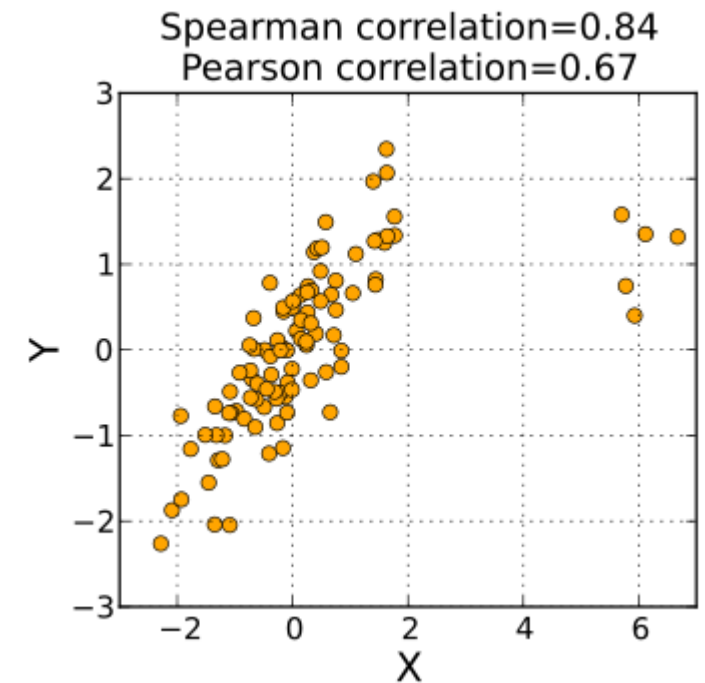
Korelační koeficient



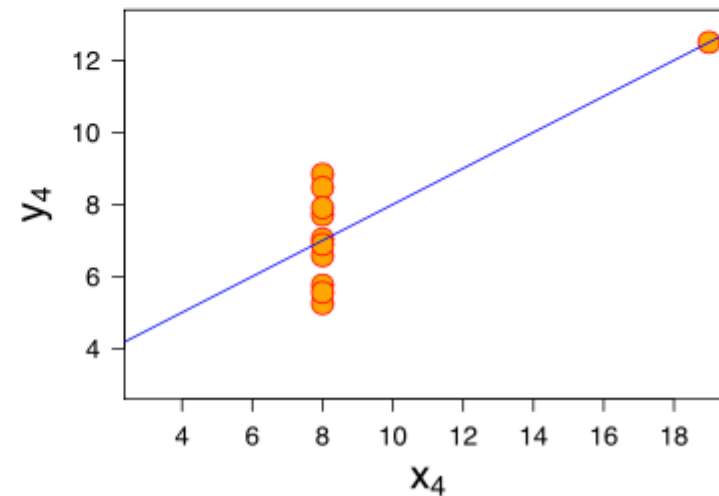
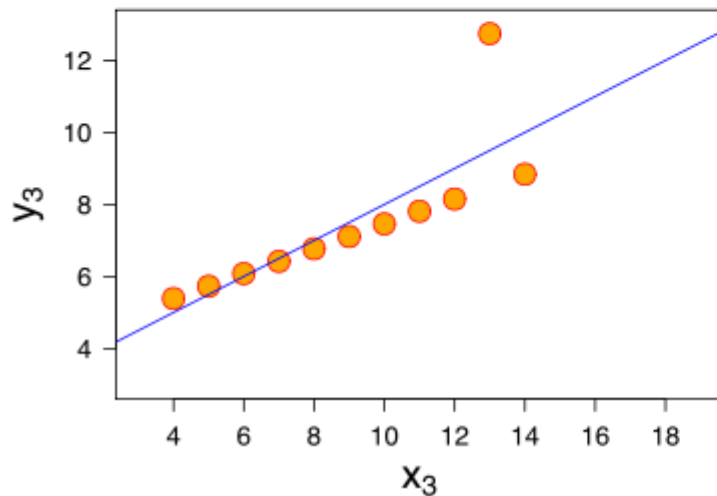
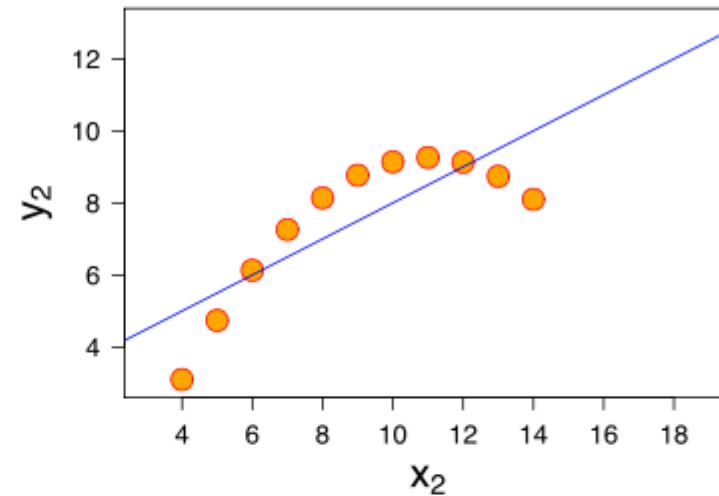
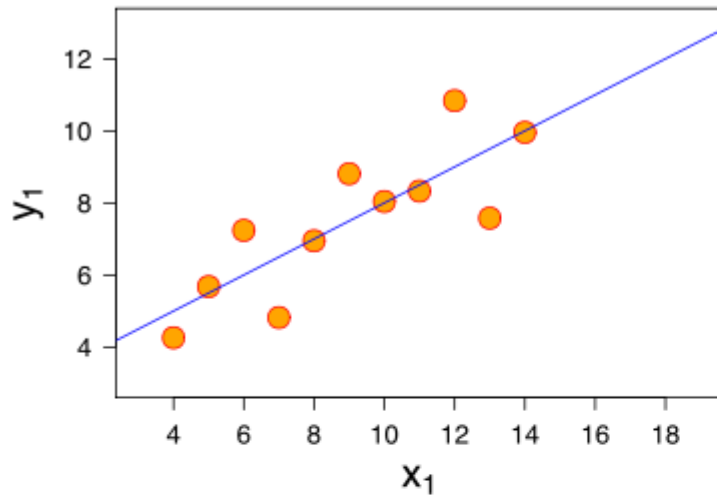
Spearmanova pořadová korelace

$$\rho = 1 - \frac{6 \sum d_i^2}{n(n^2 - 1)}$$

$$\rho = \frac{\sum_i (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_i (x_i - \bar{x})^2 \sum_i (y_i - \bar{y})^2}}$$



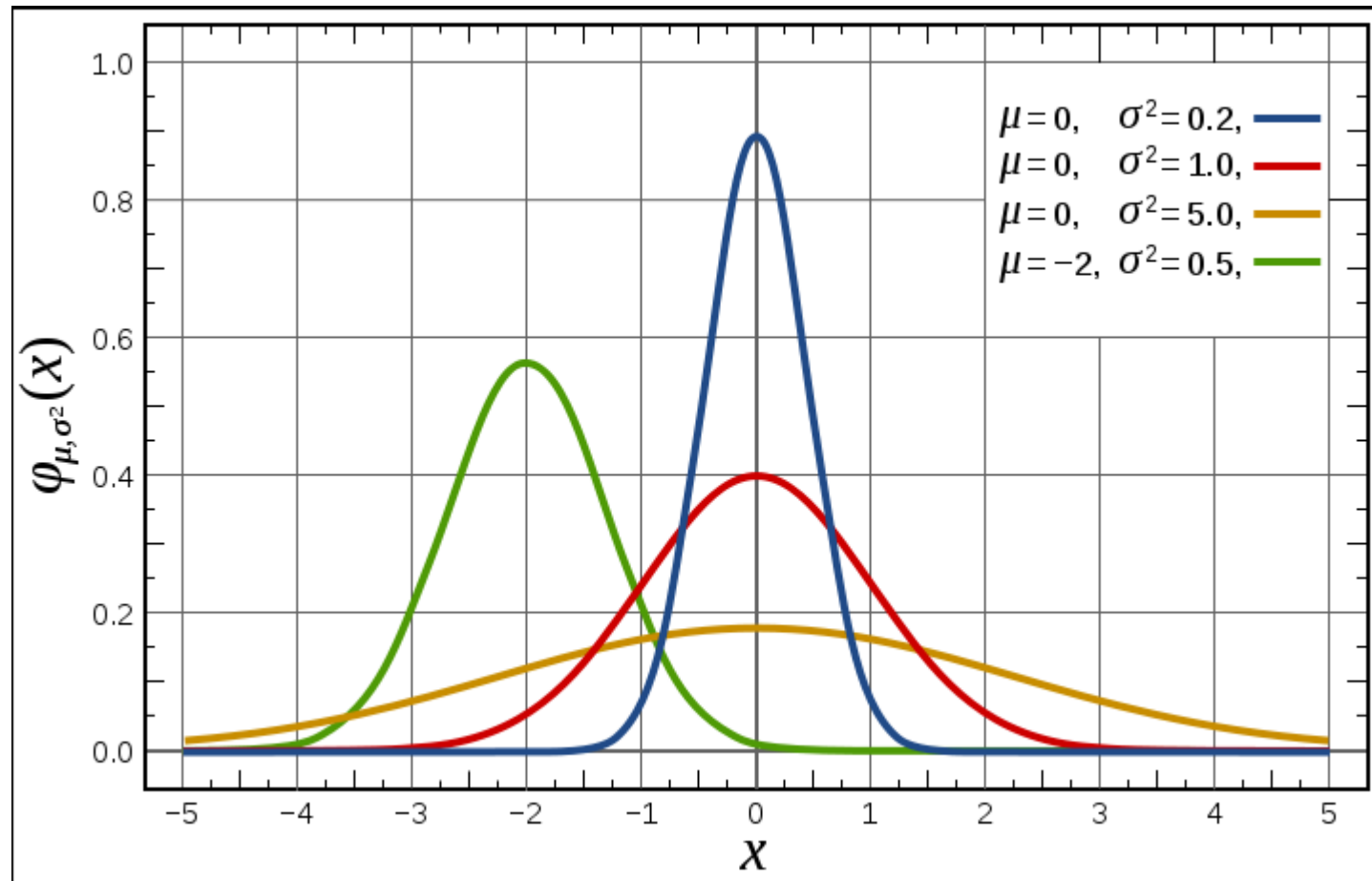
Stejný koeficient determinace



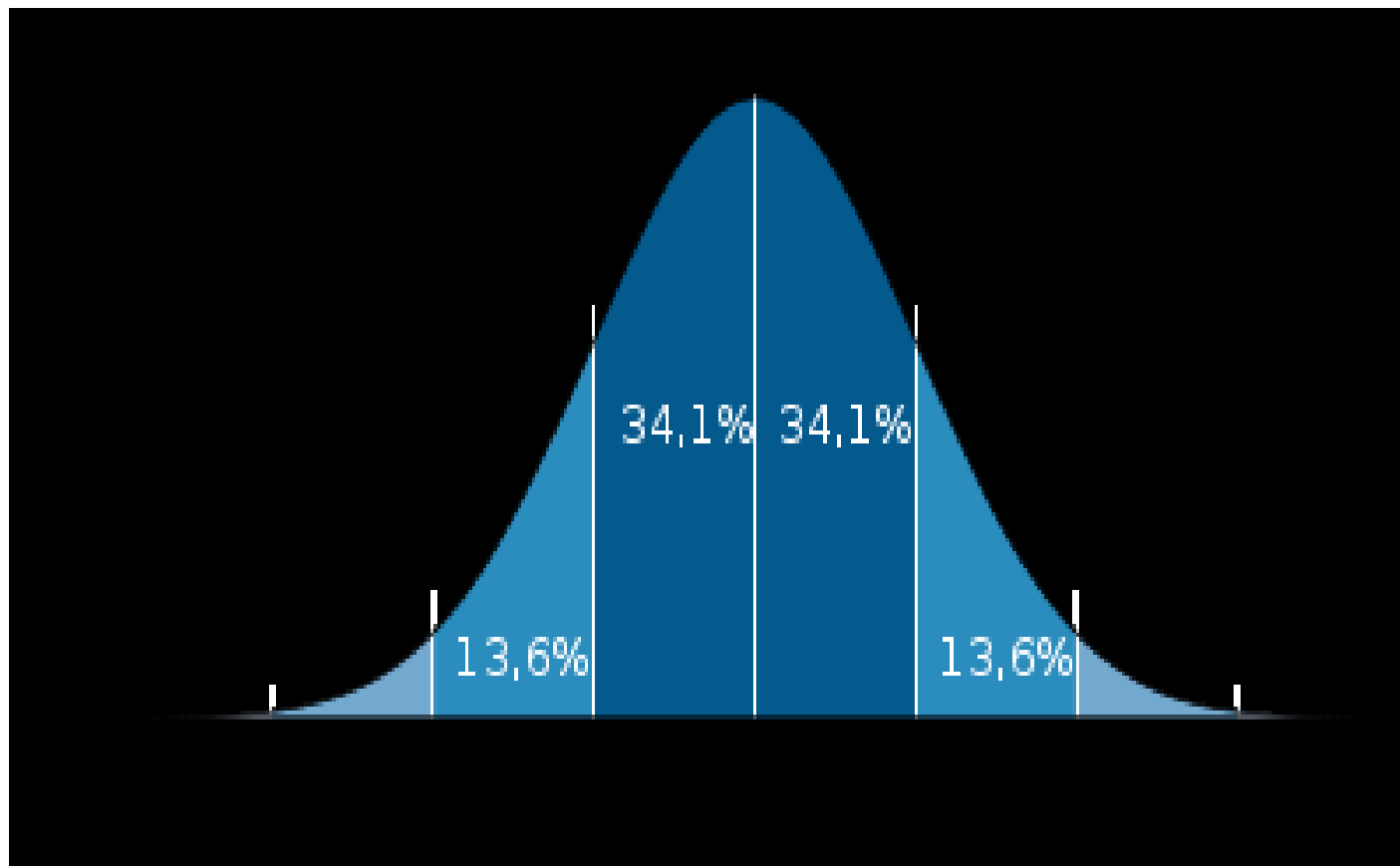
Statist - menu

- Testy rozložení dat (volba 3)
 - 0 = Main menu
 - 1 = t-test for comparison of two means of two samples
 - 2 = t-test for comparison of pairwise ascertained samples
 - 3 = Test of normal distribution (KS-Lilliefors-Test)
 - 4 = Chi²-fourfold-test
 - 5 = Chi² two-items-test
 - 6 = u-test (Test of independence of two samples)
 - 7 = H-test (Kruskal-Wallis) for k independent samples
 - 8 = Wilcoxon-Rank-test for pairwise ascertained samples
 - 9 = Chi²-test of equal frequency
 - 10 = Chi²-test of correspondence between measured and theoretical frequency

Normální rozložení



Normální rozložení



Statist - menu

- Různé (volba 4)
 - 0 = Main menu
 - 1 = Standard deviation, mean, median, etc.
 - 2 = Probit analysis
 - 3 = Outliers & Box-Whisker-plot
 - 4 = Percentiles
 - 5 = Frequency table
 - 6 = Compare means
 - 9 = Enter gnuplot commands

Statist - menu

- Manipulace s daty (volba 5)
 - 0 = Main menu
 - 1 = Log-transformation (base 10)
 - 2 = Invert values ($1/x$)
 - 3 = z-transformation $[(x-\mu)/\sigma]$
 - 4 = Sort
 - 5 = Join columns
 - 6 = Exponentiation to base 10
 - 7 = Create columns for weighted mean
 - 8 = Log-transformation (natural logarithm)
 - 9 = Exponentiation to base 'e'