Anova 1 factor

Menu:	QC.Expert	ANOVA

ANOVA is a common acronym for analysis of variance. The ANOVA module can be used for instance to check whether several batches of material or several material types differ with respect to a measured variable or characteristic. The groups to be compared are defined by *factor levels* (e.g. material can be a factor with levels corresponding to the individual material types). An inter-laboratory comparison might serve as another example with individual laboratories corresponding to different levels of the "Laboratory" factor. Similarly, various purity levels or various production lines can be compared with respect to a response variable. The goal is to decide, whether response variable means differ among different levels of a factor. This is achieved by testing the hypothesis about equality of means among different factor levels. The ANOVA test assumes normality within groups defined by the factor levels, variance homogeneity across the groups and independence of all observations. The procedure should be applied to data free of outliers. ANOVA results are accompanied by the z-score plot. The plot shows estimated group means together with their error bars, obtained as ±3.SEM (standard error of the mean). The z-score is commonly used for inter-laboratory comparisons or to compare potential suppliers.

Data and parameters

Scheffe's pairwise comparisons can be requested by checking *Pairwise comparisons* within the Analysis type part of the ANOVA dialog panel. In case that the overall ANOVA test is significant and some group is found to be significantly different by the Scheffe's test, an additional analysis is recommended. The ANOVA and Scheffe's tests should be recomputed, omitting the observations corresponding to the differing group. It can happen that the variance is reduced and the second analysis finds a previously undetected difference. This procedure can be repeated until no significant differences are found. The z-score can be requested by checking the appropriate selection in the ANOVA dialog panel (see Figure 18). When no values are entered in the *Center* and *Standard deviation* fields, the z-score plot uses grand mean and residual standard deviation respectively. *Significance level* sets the significance level for all tests, 0.05 is a commonly used value. Data can be entered in two formats specified by either *Column wise* or *By factor* button.

Column wise format:
Data corresponding to different levels of a factor are entered in different columns. Number of data points in different columns/groups can be different. The minimum column number (i.e. the number of factor levels) is 2. The minimum row number is 2. Column names should correspond to levels of the factor, e.g. Line A, Line B, Line C. Data columns can be selected in the *Columns* field of the ANOVA dialog panel, see Figure 18. The *Select all* button selects all columns of a given data sheet for the ANOVA analysis. All columns of the current data sheet are selected by default.
Data example

External	Lab 1	Lab 2	Lab 3
1.47	1.24	2.32	3.6
1.75	0.94	2.4	7.3
1.09	1.84	1.45	2.65
3.09	0.3	1.86	8.2

By factor arrangement*:*
Data are entered in two columns. The column chosen in the *Factor* field contains level codes. The column chosen in the Data column contains values of the response variable.
Data example:

Origin	Quality
UKR	17.17
GER	23.73

GER	23.7
ARG	24.78
BRA	27.91
SWE	23.19
BRA	26.87
ARG	24.59
UKR	19.5

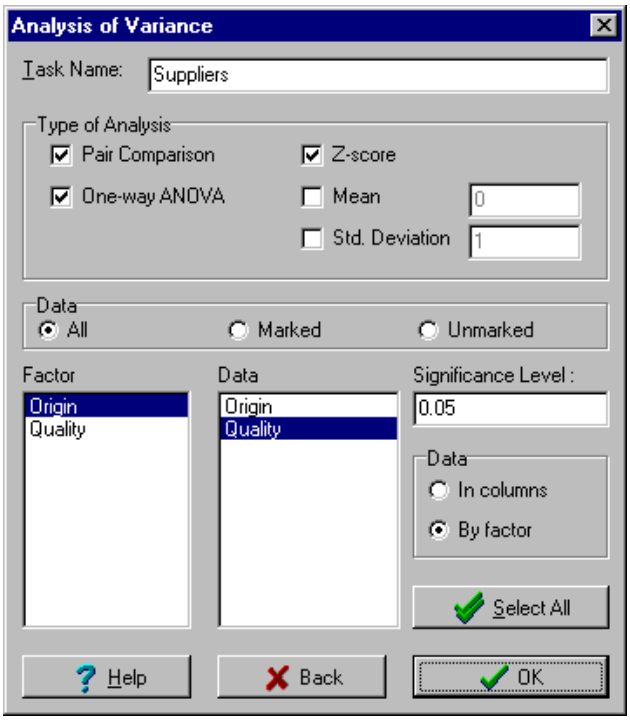


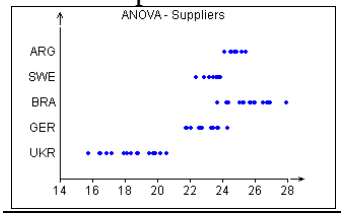
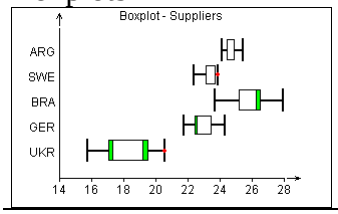
Fig. 1 ANOVA dialog panel

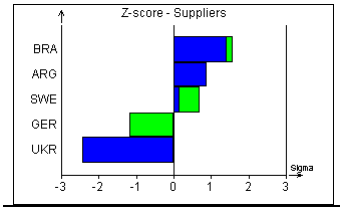
Protocol

Analysis of variance - ANOVA	
Number of levels Column	Number of factor levels.
Sample size	Number of valid data points within a factor level.
Level effects	Mean effect for each level of the factor. Difference between the group mean and the grand mean.
Means	Group means for each group defined by factor level.
Total mean	Overall mean, computed from all data.
Total variance	Variance computed from all data, regardless the grouping given by the factor levels. This is a legitimate variance estimate only if the factor has no effect and differences between levels are purely random.
Total mean square (corrected)	Mean of squared differences between observations and the grand mean.
Residual variance	Within group variance estimate. It is a legitimate variance estimate even if there

	are differences among response means for different factor levels.
Residual sum of squares	Sum of squared differences between observations and their respective group means.
Total sum of squares (corrected)	Sum of squared differences between observations and the grand mean..
Explained sum of squares	Difference between (corrected) total and residual sum of squares. It corresponds to the variability, explained by the differences among means for different factor levels.
Conclusion	Conclusion of the ANOVA test, describing in words whether the factor influences the response variable.
Factor	Indicates whether the factor has a significant effect.
Theoretical	Critical value, corresponding to the significance level chosen in the ANOVA dialog panel.
Calculated	Calculated value of the test criterion.
p-value	The smallest significance level for which the hypothesis of no factor effect is rejected using the observed data. When the p-value is smaller than a specified significance level, the factor is statistically significant.
Pairwise comparisons, Scheffe's method	All pairwise comparisons.
Pair	
Difference	Difference between group means, corresponding to different factor levels.
Conclusion	Conclusion of the test, indicating in words whether the factor is significant.
p-value	The smallest significance level for which the hypothesis of no difference among the means is rejected using the observed data. When the p-value is smaller than a specified significance level, the factor is statistically significant.
Z-score	Comparison of factor levels in terms of the standardized response group means.
Standardized value	Standardized value. A value smaller than –3 or larger than 3 might indicate that the level is different from other levels.
95% interval	Half of the 95% confidence interval length.
Difference	Limit of the 95% confidence interval.

Graphs

ANOVA plot 	ANOVA plot shows individual data values versus factor levels, so that it is possible to check whether the groups defined by the factor levels differ in mean or variance. Observed values are plotted along x-axis, factor levels are plotted along y-axis. Plot inspection, together with the *Basic data analysis* module results can reveal outliers.
Boxplots 	The boxplots are useful when comparing response among factor levels. Plot inspection, together with the *Basic data analysis* module results can reveal outliers.
Z-score plot	Comparison of factor levels in terms of the standardized response group means. The reference values is 0. The vertical lines correspond to the ±3SEM (standard error of the mean) limits. Shorter bar corresponds to the standardized value. The groups having values smaller than –3 or larger than 3



are considered to be different from other groups. Longer bar corresponds to limit of the 95% confidence interval.