DESIGN OF EXPERIMENT
&software package CHEMOFACE

Dr. Miroslava Nedyalkova

Classical statistics

- Experimental errors – random and systematic;
- Normal distribution of experimental results – the backbone of classical statistics;
- Statistical tests for normal distribution: F – test, Student's t – test etc – use and significance;
- Analytical metrology based on classical statistics – uncertainty, sensitivity, limit of detection, selectivity, reliability of the analytical signal, calibration

CORRELATION ANALYSIS

- Correlation relationship
- How to determine and interpret correlation
- Correlation coefficient
- Significance of correlation coefficient
- Misinterpretation of correlation

How to calculate correlation coefficient ?

- *n* – number of observations for groups of results X and Y

$$r = \frac{n\sum_{i=1}^{n} x_i y_i - \sum_{i=1}^{n} x_i \sum_{i=1}^{n} y_i}{\sqrt{\left[n\sum_{i=1}^{n} x_i^2 - \left(\sum_{i=1}^{n} x_i\right)^2 \right]\left[n\sum_{i=1}^{n} y_i^2 - \left(\sum_{i=1}^{n} y_i\right)^2 \right]}}$$

Correlation coefficient values

The possible values of the correlation coefficient are

$$-1 \leq r \leq +1$$

At r=1 and *r= –1 one could expect linear dependence between the compared groups of results*

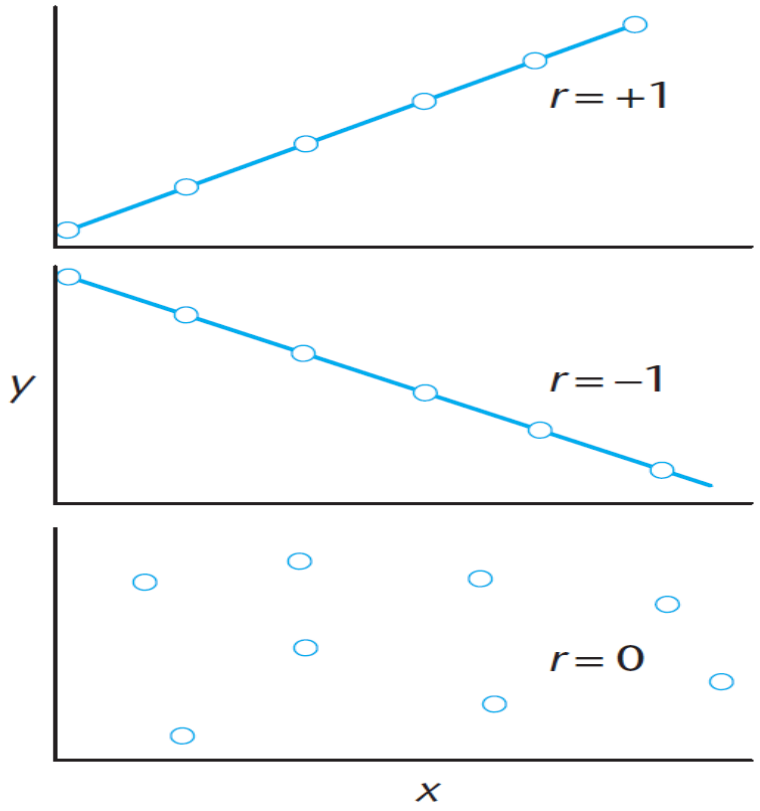
At r=0 there is no linear correlation but it does not exclude more complex relationship.

Thus, *r* assesses only the level of relationship with respect to linearity

How to interpret values of r between + 1 and - 1

- It is possible to introduce something as ***a rank*** of the correlation coefficient depending on its absolute value.
- **High correlation: 0.75 – 1.00**
- **Significant correlation: 0.50 – 0.75**
- **Low correlation: 0.25 – 0.50**
- **Insignificant correlation: 0.10 – 0.25**
- **Lack of correlation: less than 0.1**

Positive, negative and lack of correlation



Important conclusion

- Correlation does not mean casual link between the groups of interest

Reasons for jokes with statistics

- Danish rural life – white stork nests (number of objects X) and newborn kids (number of objects Y)
- Weight (or haunch) dimension and intelligence quatation
- Blondies and IQ

Significance of correlation coefficient

- The test should indicate if the correlation coefficient is statistically different from zero:

$$t = \frac{r}{\sqrt{1 - r^2}} \sqrt{n - 2}$$