C. Roussel

Physical and Chemical Analyses of Materials

Chemical analysis: data analysis

Introduction

⇨ In chemical analysis one has to perform several determinations to obtain a good overview of the chemical composition of a given sample.

⇨ Therefore, a chemical determination is expressed as a mean with a variation around this mean, the standard deviation **SD**.

⇨ The chemical determination can be related to a molar concentration, a mass concentration, a mass percentage, an atomic percentage…

⇨ For a value **x** determined with sufficient assays, **N > 20**, one can qualify the mean μ_x as a function of x_0, the true value of **x** and the standard deviation σ_x:

$$\mu_x - \sigma_x \leq x_0 \leq \mu_x + \sigma_x$$

⇨ The analytical determination is then: $\mu_x \pm \sigma_x$

⇨ One can see two types of errors that degrade the quality of the analysis. The difference between the measured and the true value ($\mu_x - x_0$) and the dispersion of the different determinations around the mean value ($\pm \sigma_x$).

Quality of analysis

Introduction

⇨ The quality of an analysis is defined in terms of precision, trueness and sensibility.

↳ The precision is linked to aleatory errors.

↳ The trueness is linked to systematic errors.

↳ The sensibility stems from the quality of the analytical device. The sensibility is qualified through the limit of detection (**LOD**) and the limit of quantification (**LOQ**) of an analyte.

$$\textbf{Analysis quality} \propto \begin{cases} \textbf{trueness (systematic errors)} \\ \textbf{precision (aleatory errors)} \\ \textbf{sensibility (LOD and LOQ)} \end{cases}$$

General features

⇨ In general, an analytical result stems from several determinations.

⇨ These determinations are expressed as a mean with a variation around this mean, the standard deviation.

⇨ The standard deviation represents the precision of the measurement.

⇨ From a mathematic point of view, one has to consider two different domains according to the number of determinations **N** performed to provide a mean measurement.

↳ For **N** ≤ **20**, the set of determinations represent a sample which follows a Binomial statistic law, a discrete distribution.

↳ For **N** > **20**, the set of determinations represent a population which follows a Gaussian statistic law, a continuous distribution.

⇨ In chemical analysis the goal is to study a sample which is representative enough to provide a precise information on the population from which the sample is picked up.

↳ To ensure an accurate analysis, one has to analyse different samples to get a representative information on a whole population, *e.g.* the analysis of the mercury content in the Leman Lake.

Mean

⇨ The mean corresponding to the determination of a value **x** is calculated as the arithmetic mean and is noted $\bar{x}$ for a sample (**N ≤ 20**) and μ_x for a population (**N > 20**):

$$\bar{x} \text{ ou } \mu_x = \frac{\sum_{i=1}^{N} x_i}{N}$$

Precision

⇨ The precision is expressed by the standard deviation **SD**, the relative standard deviation **RSD** or the coefficient of variation **CV**.

⇨ The **SD** is always associated to a mean and it represents the dispersion of the determinations around this mean. The lower dispersion, the better precision.

⇨ The **SD** accounts for the aleatory errors observed during the analysis which can't be corrected. They are oscillating around the mean with the same positive and negative magnitudes and are noted as ± **SD**.

⇨ The **SD** associated to the determinations is noted s_x for a sample (**N ≤ 20**) and σ_x for a population (**N > 20**).

⇨ The **SD** is given by:

$$s_x = \sqrt{\frac{\sum_{i=1}^{N}\left(x_i - \bar{x}\right)^2}{N-1}} \quad \text{and} \quad \sigma_x = \sqrt{\frac{\sum_{i=1}^{N}\left(x_i - \mu_x\right)^2}{N}}$$

↳ In these expressions, **N** represents the number of determinations which is also called the degrees of freedom.

⇨ The **RSD** is:

$$s_{rx} = \frac{s_x}{\bar{x}} \quad \text{and} \quad \sigma_{rx} = \frac{\sigma_x}{\mu_x}$$

⇨ The **CV** is:

$$CV_x = \frac{s_x}{\bar{x}} \times 100 \quad \text{and} \quad CV_x = \frac{\sigma_x}{\mu_x} \times 100$$

⇨ The analytical determination of a value **x** from a sample, expressed in terms of **SD** is then: $\bar{x} \pm s_x$

Trueness or bias

⇨ The trueness or bias relates to the systematic errors arising from *e.g.* a calibration mistake(s).

⇨ In contrast with aleatory errors, systematic errors can be and must be corrected when calibrating the analytical device(s).

⇨ Moreover, these errors are either positive or negative that leads to either an over- or an under-estimation of a given value **x** respectively.

⇨ These errors are not associated to the number of determination performed to get the mean of the value **x**.

⇨ It can be expressed as the absolute difference between the measured mean of **x** and the true value of **x** namely $\mathbf{x_0}$:

$$\Delta\bar{\mathbf{x}}(\mathbf{x}_0) = \bar{\mathbf{x}} - \mathbf{x}_0 \quad \textbf{and} \quad \Delta\mu_{\mathbf{x}}(\mathbf{x}_0) = \mu_{\mathbf{x}} - \mathbf{x}_0$$

⇨ It can be also expressed in percentage as:

$$\mathbf{E_r}\bar{\mathbf{x}}(\mathbf{x}_0) = \frac{|\Delta\bar{\mathbf{x}}(\mathbf{x}_0)|}{\mathbf{x}_0} \times \mathbf{100} \quad \textbf{and} \quad \mathbf{E_r}\mu_{\mathbf{x}}(\mathbf{x}_0) = \frac{|\Delta\mu_{\mathbf{x}}(\mathbf{x}_0)|}{\mathbf{x}_0} \times \mathbf{100}$$

Accuracy

⇨ The accuracy, incorporates both types of error.

↳ In general, when the systematic errors are eliminated by calibration, the accuracy is only linked to the **SD** (see calibration). As soon as the systematic is lower than the **SD**, the bias should not be corrected.

Repeatability

⇨ The repeatability describes the precision of within-run replicates.

Reproducibility

⇨ The reproducibility describes the precision of between-run replicates.

Repeatability versus reproducibility

⇨ The reproducibility of a method is normally expected to be poorer (*i.e.* with larger random errors) than its repeatability.

Error propagation

⇨ When the expressed **y** value is not a measured one but stems from a combination of measured values **a**, **b** and **c**, then the reported error on **y** must take into account the errors on the combined values **a**, **b** and **c**. The table below presents serval combinations when the determination is made on a sample (**N ≤ 20**):

operation	Precision*
$\bar{y} = \bar{a} + \bar{b} - \bar{c}$	$s_y = \sqrt{s_a^2 + s_b^2 + s_c^2}$
$\bar{y} = k_a\bar{a} + k_b\bar{b} - k_c\bar{c}$	$s_y = \sqrt{(k_a s_a)^2 + (k_b s_b)^2 + (k_c s_c)^2}$
$\bar{y} = \left(\bar{a} \times \bar{b}\right)/\bar{c}$	$\frac{s_y}{\bar{y}} = \sqrt{\left(\frac{s_a}{\bar{a}}\right)^2 + \left(\frac{s_b}{\bar{b}}\right)^2 + \left(\frac{s_c}{\bar{c}}\right)^2}$
$\bar{y} = f(\bar{x})$	$s_y = s_x \left\lvert\frac{d\bar{y}}{d\bar{x}}\right\rvert$
$\bar{G} = f(\bar{x}, \bar{y}, \bar{z})$	$s_G = \sqrt{\left(\frac{\partial f(\bar{x},\bar{y},\bar{z})}{\partial \bar{x}}\right)^2 s_x^2 + \left(\frac{\partial f(\bar{x},\bar{y},\bar{z})}{\partial \bar{y}}\right)^2 s_y^2 + \left(\frac{\partial f(\bar{x},\bar{y},\bar{z})}{\partial \bar{z}}\right)^2 s_z^2}$

*The formulas are still identical for a population (**N > 20**), only the manner to calculate the standard deviation must be appropriate.

⇨ The table below presents serval combinations when the determination is made on a sample ($\mathbf{N \leq 20}$):

operation	trueness*
$\bar{y} = \bar{a} + \bar{b} - \bar{c}$	$\Delta y = \Delta a + \Delta b + \Delta c$
$\bar{y} = k_a\bar{a} + k_b\bar{b} - k_c\bar{c}$	$\Delta y = k_a\Delta a + k_b\Delta b + k_c\Delta c$
$\bar{y} = \left(\bar{a} \times \bar{b}\right)/\bar{c}$	$\frac{\Delta y}{\bar{y}} = \left(\frac{\Delta a}{\bar{a}}\right)+\left(\frac{\Delta b}{\bar{b}}\right)+\left(\frac{\Delta c}{\bar{c}}\right)$
$\bar{y} = f(\bar{x})$	$\Delta y = \Delta x \left\lvert\frac{d\bar{y}}{d\bar{x}}\right\rvert$
$\bar{G} = f(\bar{x}, \bar{y}, \bar{z})$	$\Delta_G = \left\lvert\frac{\partial f(\bar{x},\bar{y},\bar{z})}{\partial \bar{x}}\right\rvert \Delta x + \left\lvert\frac{\partial f(\bar{x},\bar{y},\bar{z})}{\partial \bar{y}}\right\rvert \Delta y + \left\lvert\frac{\partial f(\bar{x},\bar{y},\bar{z})}{\partial \bar{z}}\right\rvert \Delta z$

*The formulas are still identical for a population ($\mathbf{N > 20}$).

Instrument calibration

General features

⇨ In chemical analysis, the calibration of the analytical device is often performed in the linear domain of the detector.

⇨ In this case, the calibration method is based on a linear regression.

⇨ The linear regression can be either a simple or a weighted linear regression depending on the considered situation.

Calibration through a simple linear regression

⇨ The simple regression is based on the mean least squares method. For a linear equation $\mathbf{y = mx + b}$, the method is based on the minimization of the distance between the line and the real point along the x-axis or along the y-axis or along both.

⇨ In general, the x-axis represents the values of the standards. We consider that the difference observed between the optimal and the real points stems from the instrument and is reported on the y-axis corresponding to the response, that is either the signal $\mathbf{S}$ ($\mathbf{y = S}$) or the integrated signal ($\mathbf{y} = \int \boldsymbol{Sdx}$) of the instrument.

↳ In this case we attribute all the errors to the instrument measurement.

↳ The regression method will then minimize the distance between the regression line and the measured points along the y-axis.

⇨ To perform this linear regression using **N** standards, one has to calculate the following values:

$$\mathbf{s_{xx} = \sum_{i=1}^{N} x_i^{\,2} - N\left(\overline{x}\right)^2 \quad and \quad s_{yy} = \sum_{i=1}^{N} y_i^{\,2} - N\left(\overline{y}\right)^2 \quad and \quad s_{xy} = \sum_{i=1}^{N} x_i y_i - N\left(\overline{x}\right)\left(\overline{y}\right)}$$

⇨ The coefficient of correlation **r** which is related to the intensity of the coupling between **x** and **y** through a line is:

$$\mathbf{r = \frac{\sum_{i=1}^{N}\left(x_i - \overline{x}\right)\left(y_i - \overline{y}\right)}{\sqrt{\sum_{i=1}^{N}\left(x_i - \overline{x}\right)^2 \times \sum_{i=1}^{N}\left(y_i - \overline{y}\right)^2}} = \frac{s_{xy}}{\sqrt{s_{xx}s_{yy}}}}$$

↳ For a perfect line of increasing slope: **r = 1**, for a perfect line of decreasing slope: **r = - 1** and for a line of slope equal to zero, *i.e.* a perfect line parallel to the x-axis: **r = 0**.

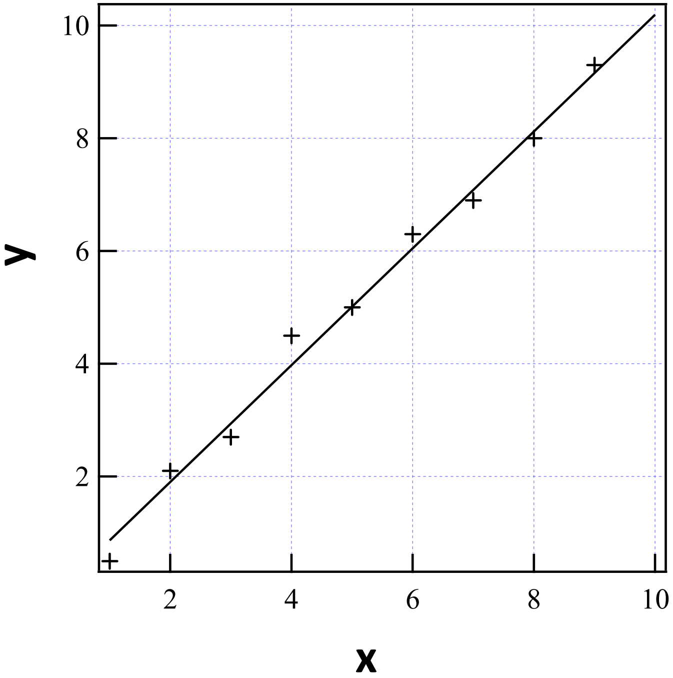
- For a linear regression, the goodness of fit (GOF) $\mathbf{R^2}$ is $\mathbf{R^2 = r^2}$.

- The regression parameters **m**, **b**, $\mathbf{s_m}$ and $\mathbf{s_b}$ for an equation $\mathbf{y = m(\pm s_m)x + b(\pm s_b)}$ are the followings:

$$m = \frac{s_{xy}}{s_{xx}} \quad \text{and} \quad b = \bar{y} - m\bar{x}$$

$$s_m = \sqrt{\frac{s_r^{\,2}}{s_{xx}}} \quad \text{and} \quad s_b = s_r\sqrt{\frac{s_{xx} + N\left(\bar{x}\right)^2}{N s_{xx}}}$$

$$S_r = \sqrt{\frac{S_{yy} - m^2 S_{xx}}{N-2}}$$



- A simple linear regression is performed when the **SD** of each standard used for the calibration are rather similar which is often the case.

Limits of detection and of quantification (LOD and LOQ)

⇨ In general, the **LOD** and **LOQ** which constitute the sensibility of an analytical method can be obtained through the linear regression performed to calibrate the instrument.

⇨ The calibration domain is delimited by the lower limit of quantification, also called limit of quantification **LOQ** and the upper limit of quantification **ULOQ**.

↳ Between **LOQ** and **ULOQ**, the response of the detector is linear and the values obtained are accurate.

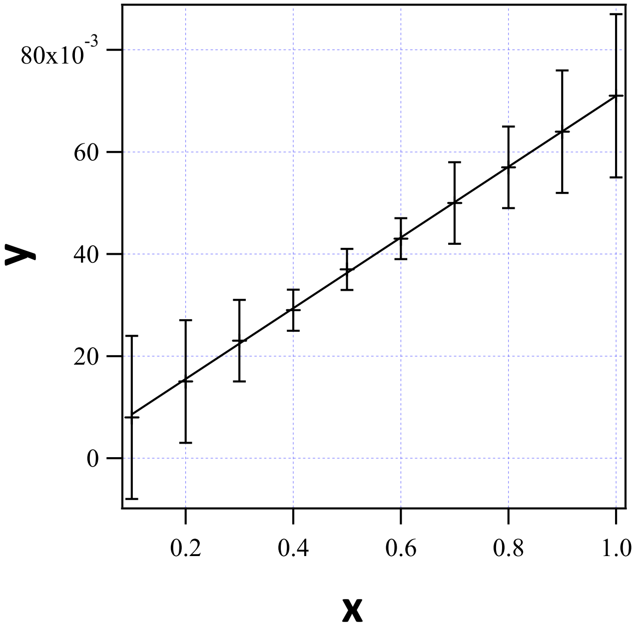
↳ Above the **ULOQ**, the response of the detector tends to reach an asymptote, the detector is saturated.

↳ Below the **LOQ**, the response of the detector is not linear, accurate values can not be obtained in this specific domain.

↳ In this specific domain the limit of detection **LOD** gives the minimal value below which one can not claim that the analyte is detected with this analytical method on this instrument.

⇨ In addition, to obtain an accurate analysis, one has to consider the **SD** of each standard used to build the calibration as shown below.

↳ The calibration curve below includes the **SD** of each standard used for the calibration:



↳ In this graph, one can observe that the **SD** of the standards, is minimal in the middle of the linear domain and tends to increase when approaching both **LOQ** and **ULOQ**.

↳ When approaching the **ULOQ**, the **SD** increase is related to the detector saturation. This constitutes the instrumental limitation.

↳ When approaching the **LOQ**, the **SD** increase is related to the precision of the analytical glasswares used to prepare the standards.

⇨ As a conclusion, the best precision is obtained from the middle of the calibration curve.

⇨ The **LOD** and **LOQ,** considering the **SD** of the background signal $\mathbf{s_{bg}}$, are given below:

$$\mathbf{LOD = mx_{LOD} + b \quad \Rightarrow \quad y_{bg} + 3s_{bg} = mx_{LOD} + b}$$

$$\mathbf{LOQ = mx_{LOQ} + b \quad \Rightarrow \quad y_{bg} + 10s_{bg} = mx_{LOQ} + b}$$

$$\mathbf{x_{LOD} = \frac{y_{bg} + 3s_{bg} - b}{m} \quad and \quad x_{LOQ} = \frac{y_{bg} + 10s_{bg} - b}{m}}$$

⇨ When $\mathbf{y_{bg}}$ and **b** compensate, the **LOD** and **LOQ** read:

$$\mathbf{x_{LOD} = \frac{3s_{bg}}{m} \quad and \quad x_{LOQ} = \frac{10s_{bg}}{m}}$$

⇨ When the background signal and **y** are the not of the same dimension, *i.e.* a background signal as an intensity and **y** as a peak area (integration of the intensity over time) $\mathbf{s_{bg}}$ must be replaced by $\mathbf{s_r}$:

$$\mathbf{x_{LOD} = \frac{3s_r}{m} \quad and \quad x_{LOQ} = \frac{10s_r}{m}}$$

⇨ The proposed **LOD** and **LOQ** must be checked by building standards at the determined **LOD** and **LOQ** which are tested at least 10 times.

↳ These values will be valid if during this test, the proposed values of **LOD** and **LOQ** are found at 95%.

Bias of a calibration

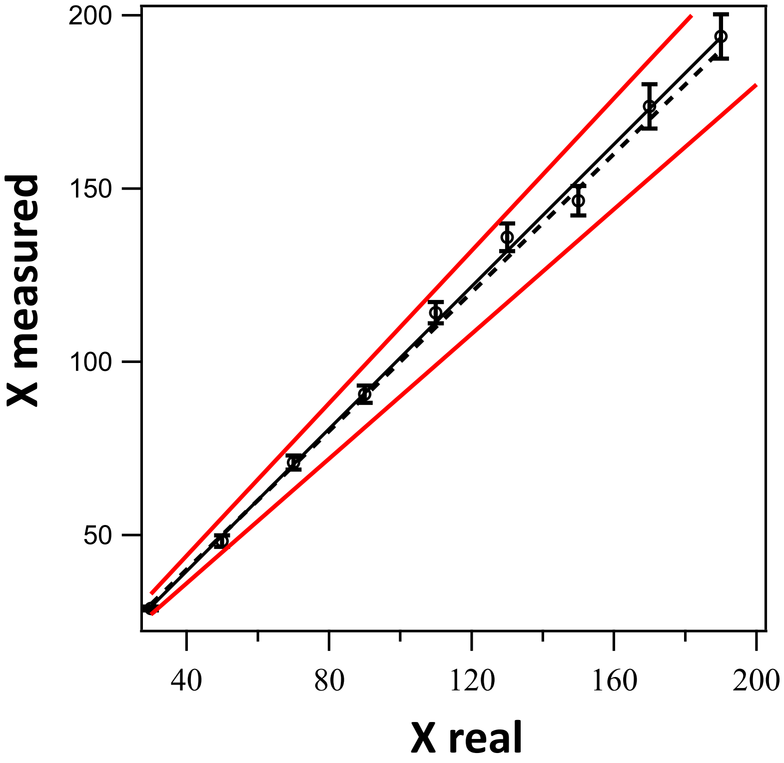
⇨ To determine the bias of a calibration, one has to build some standards different from the one used to build the calibration curve but included in the calibration domain.

↳ Each standard is measured and determined according to the linear regression parameters of the calibration curve.

↳ Then, a curve expressing the measured value as a function of the true value is plotted.

↳ This curve helps in determining the bias of an analytical method. If there is no bias, this curve must be the first bisector, *i.e.* $\mathbf{y} = \mathbf{x}$.

⇨ Below is shown a validation curve:



↳ The solid black line represents the first bisector ($\mathbf{x_{measured}} = \mathbf{x_{real}}$).

↳ The dashed black line represents the validation curve.

↳ The two red solid lines represent the limits of confidence of 90% ($\mathbf{x_{measured}} = \mathbf{0.9\ x_{real}}$ and $\mathbf{x_{measured}} = \mathbf{1.1\ x_{real}}$).

↳ At each standard is associated its **SD**.

⇨ One can observe that the slope of the validation curve is a slightly lower than the slope of the first bisector thus attesting from a tiny bias.

⇨ As the bias is contained in the **SD** for the all points, thus the bias is negligible so the validation domain cover the whole linearity domain.

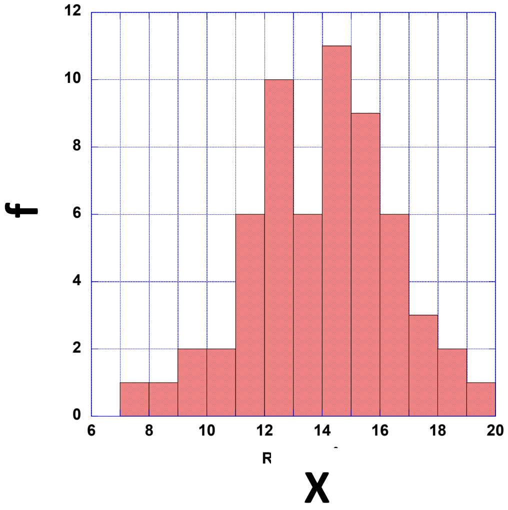
Sampling

General features

⇨ To consider a population, the number of determinations must be greater than 20. There is no real interest in performing 20 determinations on the same sample with the same instrument.

⇨ Actually, the best is to divide the initial sample into 5 subsamples and each subsample is tested at least 5 times. In this case, the analysis uses 5 different instruments.

⇨ The main problem remains into the assembly of all the data. Normally, each subsample must give the same mean value. An example is given below:



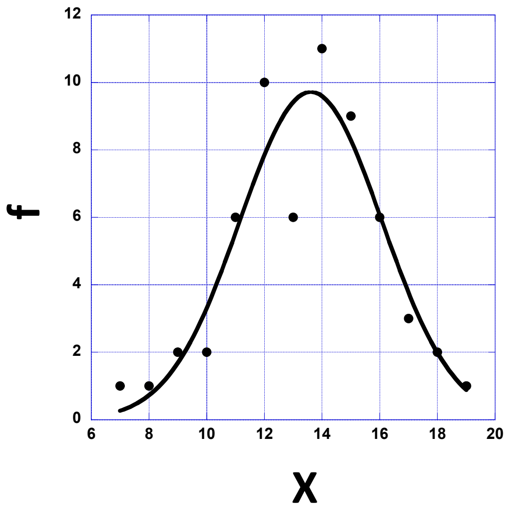
↳ On the left side is represented the frequency of an observed value as a function of the different observed values.

↳ Let says that this histogram results from the division of a sample into two subsamples tested several times.

⇨ The easiest way is two consider that all the determinations were ran under the same conditions and that the results are comparable.

↳ In this case, one can transform this discrete distribution into a Gaussian one in order to get the mean of the population and its **SD**.

⇨ Below is the fitting of the discrete distribution by a Gaussian law:



↳ The fitting leads to a mean of the population of 13.6 and an **SD** of 2.46.

↳ One of the problem of this fitting is that the values found with the maximum frequencies are not included in the curve.

↳ Actually, if one looks carefully at the histogram, two distributions are clearly visible. This means that at least one instrument was not correctly calibrated.

⇨ To figure out if the data coming the determinations **1** and **2** leading to $\overline{\mathbf{x_1}}$ and $\overline{\mathbf{x_2}}$ can be associated, a student test can be performed.

⇨ This test consists in calculating a student coefficient **t** (absolute value) using the data and one has to compare it to the student coefficient given in the tables:

$$\left|\mathbf{t}\right| = \frac{\left(\overline{\mathbf{x}}_1 - \overline{\mathbf{x}}_2\right)}{\mathbf{s}_{\mathbf{grx}}\sqrt{\dfrac{\mathbf{N}_1 + \mathbf{N}_2}{\mathbf{N}_1\mathbf{N}_2}}}$$

↳ In this equation, $\mathbf{N_1}$ and $\mathbf{N_2}$ are the number of tests ran for the determinations **1** and **2**.

↳ $\mathbf{s_{grx}}$ is given by:

$$\mathbf{s}_{\mathbf{grx}}\sqrt{\frac{\sum_{i=1}^{i=N_1}\left(\mathbf{x}_i - \overline{\mathbf{x}}_1\right)^2 + \sum_{j=1}^{j=N_2}\left(\mathbf{x}_j - \overline{\mathbf{x}}_2\right)^2 + \sum_{k=1}^{k=N_3}\left(\mathbf{x}_k - \overline{\mathbf{x}}_3\right)^2 + \ldots}{\mathbf{N}_1 + \mathbf{N}_2 + \mathbf{N}_3 + \ldots - \mathbf{N}_t}}$$

➪ If |**t**| < $\mathbf{t_{crit}}$ **(P = 0.05)**, then the two means are comparable, so the two series can be associated in a single one. Below is the table of student. To determine $\mathbf{t_{crit}}$ one has to refer to the following table in which the degrees of freedom, $\mathbf{N = (N_1+N_2)-2}$:

N	**LC (90%)** **P=0.10**	**LC (95%)** **P=0.05**	**LC (98%)** **P=0.02**	**LC (99%)** **P=0.01**
1	6.31	12.71	31.82	63.66
2	2.92	4.30	6.96	9.92
3	2.35	3.18	4.54	5.84
4	2.13	2.78	3.75	4.60
5	2.02	2.57	3.36	4.03
6	1.94	2.45	3.14	3.71
7	1.89	2.36	3.00	3.50
8	1.86	2.31	2.90	3.36
9	1.83	2.26	2.82	3.25

⇨ As soon as this criteria is met, one has to calculate the intra-variability and the extra-variability and to combine them to obtain the overall **SD** (**s_x**).

↳ the overall **SD** is grasped through the variances **Var** associated to the different measurements .

↳ When performing a one-way ANOVA analysis of a sample divided in **k** series, each sample being tested **n** times, one has to calculate the following parameters:

$$\textbf{mean square within-group: } \mathbf{MSW} = \frac{\mathbf{n}\sum_{\mathbf{i}} \mathbf{Var}\left(\mathbf{x_i}\right)}{\mathbf{k}\left(\mathbf{n}-\mathbf{1}\right)}$$

$$\textbf{mean square between-group: } \mathbf{MSB} = \frac{\mathbf{n}}{\mathbf{k}-\mathbf{1}}\sum_{\mathbf{i}}\left(\overline{\mathbf{x}}_{\mathbf{i}}-\overline{\mathbf{x}}\right)^2$$

$$\mathbf{Var}\left(\mathbf{W}\right) = \mathbf{MSW} \quad \textbf{and} \quad \mathbf{Var}\left(\mathbf{B}\right) = \frac{\mathbf{MSB}-\mathbf{MSW}}{\mathbf{n}}$$

$$\mathbf{s_x} = \sqrt{\mathbf{Var(B)}+\mathbf{Var(W)}}$$

↳ **s_x** is the overall **SD** to be reported when the analysis was performed on **k** different samples tested several **n** times.

Quantification

General features

⇨ In chemical analysis, quantification can be assessed by internal normalisation, external calibration, internal calibration, standard additions and standard additions using an internal standard.

⇨ Usually, the quantification is done in the validated domain of the analytical method, that is where the bias (systematic errors) is negligible, so only the **SD** is to take into account.

⇨ In the case of quantification assessed by internal normalisation, the error propagation applies to get the overall error on the result.

⇨ For external calibration, internal calibration, standard additions and standard additions using an internal standard, one has to build a calibration curve using a simple linear regression based on the mean least squares method.

Internal normalisation

⇨ In the simplest case, the determination of a given compound is its fraction among the overall compounds. This method takes into account an equivalent response of the solutes with respect to the instrument detector.

⇨ The fraction of the compound **J** is related to its response $\mathbf{y_J}$.

⇨ The fraction of **J** can be expressed in massic $\mathbf{\%_{m,J}}$ or in concentration percentage $\mathbf{\%_{m,J}}$ as:

$$\mathbf{\%_{m,J}} = 100\frac{\mathbf{y_J}}{\sum\limits_{\mathbf{J}=1}^{\mathbf{N}} \mathbf{y_J}} \quad \textbf{or} \quad \mathbf{\%_{c,J}} = 100\frac{\mathbf{y_J}}{\sum\limits_{\mathbf{J}=1}^{\mathbf{N}} \mathbf{y_J}}$$

⇨ To obtain the error on the determined fraction, one has to take care of the error on each signal and must carry them forward using the error propagation method.

External calibration

⇨ For external calibration, a series of standards of known composition is used and a linear regression of the response of each standard as a function of their values (concentration, mass, fraction…) is realised.

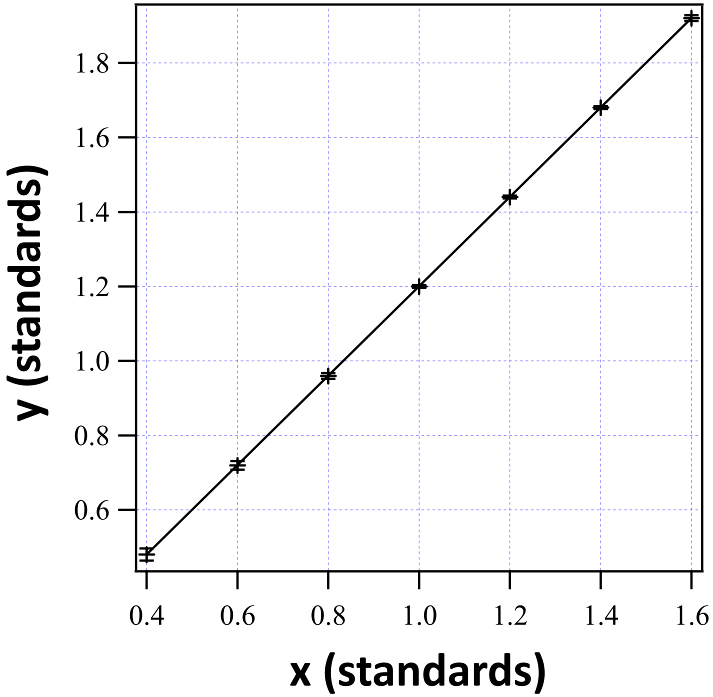
⇨ The linear regression base on the mean least squares method gives the general equation $\mathbf{y = mx + b}$ in which **y** and **x** represent the response of the standards and their values respectively.

⇨ The regression parameters **m**, **b**, $\mathbf{s_m}$ and $\mathbf{s_b}$ for an equation $\mathbf{y = m(\pm s_m)x + b(\pm s_b)}$ are the followings:

$$\mathbf{m = \frac{s_{xy}}{s_{xx}}} \quad \textbf{and} \quad \mathbf{b = \bar{y} - m\bar{x}}$$

$$\mathbf{s_m = \sqrt{\frac{s_r^{\ 2}}{s_{xx}}}} \quad \textbf{and} \quad \mathbf{s_b = s_r\sqrt{\frac{s_{xx} + N\left(\bar{x}\right)^2}{Ns_{xx}}}}$$

$$\mathbf{S_r = \sqrt{\frac{S_{yy} - m^2S_{xx}}{N-2}}}$$



⇨ The determination of an unknown sample is obtained thanks to the equation of the linear regression established with **N** calibration points measured **M** times. The **SD** of the determination based is:

$$\mathbf{s_{x_{measured}} = \frac{s_r}{m}\sqrt{\frac{1}{M} + \frac{1}{N} + \frac{\left(y_{measued} - \bar{y}\right)^2}{m^2 s_{xx}}}}$$

Internal calibration

⇨ For internal calibration, one spikes a compound **J** of known value $\mathbf{x_J}$ called a standard in the analysed sample.

⇨ Ideally, the standard should not be originally present in the analysed sample.

⇨ The equation to determine the value of **I** ($\mathbf{x_i}$) where **y** and **x** stand for the response and the value respectively is then:

$$\frac{\mathbf{y_I}}{\mathbf{y_J}} = \mathbf{k_{I/J}} \frac{\mathbf{x_I}}{\mathbf{x_J}}$$

⇨ The constant $\mathbf{k_{I/J}}$ is known as the response factor.

⇨ This constant translates the difference of response of the solutes **I** and **J** with respect to the instrument detector.

⇨ A calibration to determine the constant $\mathbf{k_{I/J}}$ must be performed using several analysed compound / standard ratios.

⇨ The linear regression base on the mean least squares method gives the general equation $\mathbf{y = mx + b}$ in which **y** and **x** represent the response ratio (probe compound / standard) and the value ratio (probed compound / standard) respectively. The slope **m** of the straight line is the constant $\mathbf{k_{I/J}}$. Knowing $\mathbf{y_I}$, $\mathbf{y_J}$, $\mathbf{k_{I/J}}$ and $\mathbf{x_J}$, one can easily determine $\mathbf{x_I}$:

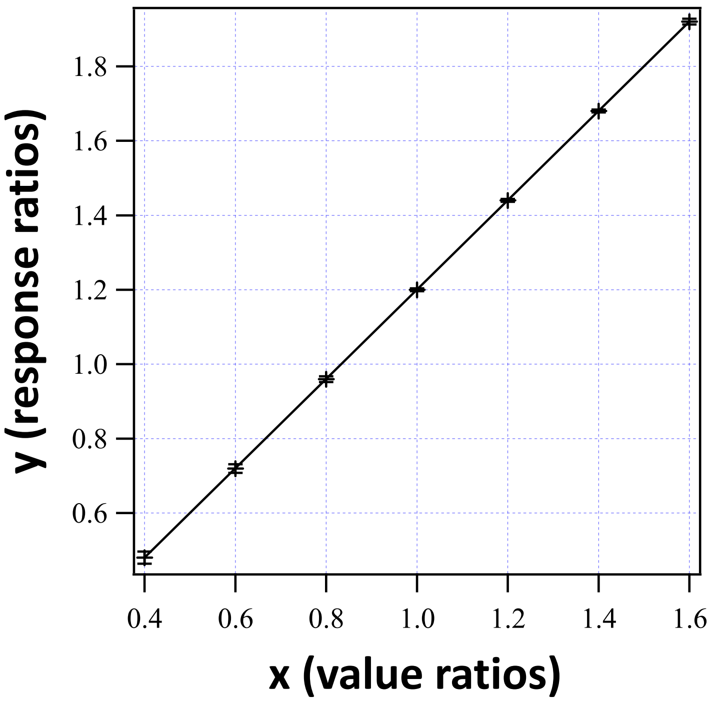
$$\frac{y_I}{y_J} = k_{I/J}\,\frac{x_I}{x_J}$$

⇨ The **SD** on the signals $\mathbf{y_I}$ and $\mathbf{y_J}$ are obtained when measuring the sample several times.

⇨ The **SD** on $\mathbf{x_J}$ depends only on the preparation of the standard (dilution, weighting…).

⇨ The **SD** on $\mathbf{k_{I/J}}$ is obtained from the linear regression ($\mathbf{s_m}$, see external calibration).

⇨ The **SD** on $\mathbf{x_I}$ is obtained by a combination of the previously mentioned errors using the error propagation method.



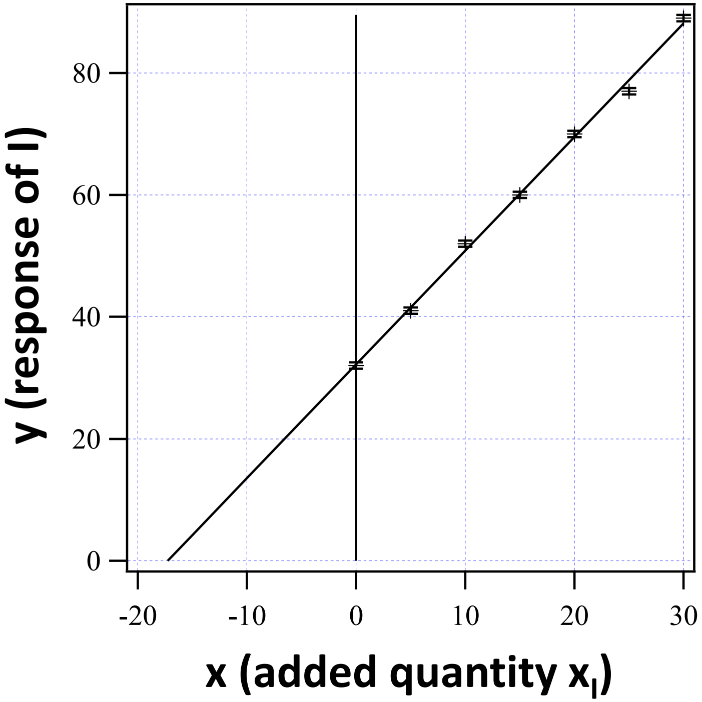
Standard additions

⇨ For standard additions, one has to spike the pure compound to determine into the analysed sample. Repeating the process several times, only the response corresponding to the compound to determine will increase.

⇨ One has to plot the response $\mathbf{y_I}$ of the probe compound **I** as a function of its spiked amount $\mathbf{x_I}$ in the analysed sample. In general, each addition corresponds to the same amount of spiked compound.

⇨ The linear regression base on the mean least squares method gives the general equation $\mathbf{y = mx + b}$ in which **y** and **x** represent the response of the probe compound and its added quantity respectively.

⇨ The determination of **I** is obtained thanks to the equation of the linear regression established with **N** standard additions and corresponds to the absolute value of the abscise corresponding to $\mathbf{y = 0}$.



⇨ The **SD** of the determination based on **N** standard additions is:

$$\mathbf{s}_{\mathbf{x}_{\mathbf{measured}}} = \frac{\mathbf{s}_{\mathbf{r}}}{\mathbf{m}}\sqrt{\frac{1}{\mathbf{N}} + \frac{\left(\overline{\mathbf{y}}\right)^2}{\mathbf{m}^2\mathbf{s}_{\mathbf{xx}}}}$$

Standard additions using an internal standard

⇨ This method combines the standard additions and the internal calibration.

⇨ The goal is to subtract to the signal of a probed compound the noise caused by a problem of repeatability of the analytical procedure.

⇨ The linear regression is then performed with the corrected response as a function of the spiked amount of the probed compound.

⇨ The amount of probed compound and its **SD** are obtained likewise the classic standard additions method.