Different approaches to estimation of measurement uncertainty in analytical chemistry

Ivo Leito
University of Tartu
ivo.leito@ut.ee

Overview

• The main question of uncertainty evaluation
• The different approaches
  – (Modelling approach)
  – Approach based on validation and QC data
• The role of method performance data
  – Precision
  – Trueness, bias
• Online course „Measurement Uncertainty Estimation in Analytical Chemistry“
The main question of uncertainty evaluation in an analytical lab:

The uncertainty sources are more or less known

There are different data available (control charts, PT results, parallel measurements …)

How to use these data to take these uncertainty sources into account?

Different approaches offer different solutions to this question

Uncertainty estimation approaches

Definition of the measurand

Single laboratory

Inter laboratory

Yes

No

Model-based?

One procedure?

Yes

No

Modelling

Component-by-component evaluation ISO GUM

Single-lab validation

Within-lab reproducibility and bias Nordtest TR537

Interlaboratory validation

Reproducibility and bias ISO 5725 ISO TS 21748

Proficiency testing (PT)

Between-lab variability ISO Guide 43 ISO 13528

Uncertainty estimates by different approaches

- Modelling (classical ISO GUM)
  - Uncertainty of an individual result of a measurement can be obtained
- Single-lab validation
  - Typical uncertainty of results obtained using a procedure in the laboratory
- Interlaboratory validation
  - Uncertainty of results obtained using the same procedure in different laboratories

These uncertainties refer to different situations!

The Modelling Approach

Component by component evaluation
**Essence**

- Based on measurement model, identification and quantification of all important uncertainty components
- Has been applied in chemistry, but often with problems

---

**Challenges with Chemical Measurements**

- Often not readily modeled
- Uncertainty contributions not readily quantified
  - Analyte losses during sample preparation
  - Interferences from other components of the sample
  - Sample inhomogeneity
  - Often insufficient information available

**Danger to underestimate uncertainty!**
Example: Moisture Content

• The model:

\[ Q_{\text{moisture}} = \frac{m_{\text{sample}} - m_{\text{sample\_after\_heating}}}{m_{\text{sample}}} \times 100\% \]

• Substituting typical balance data yields:

\[ Q_{\text{moisture}} = (12.500 \pm 0.013) \% (k = 2) \]

Example: Moisture Content

• If in addition to the balance we try to take into account:
  – Sample inhomogeneity
  – Possibly incomplete drying

• Then, substituting more realistic data, we get:

\[ Q_{\text{moisture}} = (12.50 \pm 0.88) \% (k = 2) \]
Example: Moisture Content

The difference between 0.013% and 0.88% is almost 70 times!

The intrinsic balance uncertainty sources are almost insignificant in this case!

Underestimation of uncertainty is not an “intrinsic property” of the Modelling Approach

It all depends on the implementation

But can be very work-intensive

Thus, alternative approaches have been developed
Approach Based on Validation and Quality Control Data

on the example of the Nordtest approach

http://www.nordtest.info/

Single-laboratory validation approach

Effects contributing to uncertainty

- Random
- Systematic

• The two groups of uncertainty contributions are quantified separately and then combined
**Single lab validation approach: in practice (1)**

- The main equation:

\[ u_c = \sqrt{u(R_w)^2 + u(bias)^2} \]

Within-laboratory reproducibility
This component accounts for the random effects

Uncertainty of the estimate of the laboratory and the method bias
This component accounts for the systematic effects

At "long term" level!

This and subsequent equations work with both absolute and relative values

---

**Precision component \( u(R_w) \)**

\[ u(R_w) = s_{RW} \]

is usually found from:

- the warning limits of X chart
  - using a stable control sample

- long term pooled standard deviation

Include sample preparation!

Ideally: separately for different matrices and different concentration levels!
The control sample analysis has to cover the whole analytical process

---

http://www.nordtest.info/
How to determine precision?

Example:

An analyst analysed a food sample by HPLC. He carefully homogenized the sample in a blender and took a subsample. With the subsample he carried out sample preparation (consisting of extraction, precipitation and centrifugation). As a result he obtained a clear solution. He transferred it into a 50 ml volumetric flask and filled it up to the mark with the mobile phase. He analysed 10 aliquots of this solution during the same day and calculated the within-lab reproducibility as standard deviation of the results.

Did he do it right? If not, what should he do differently?

Determining precision when sample is stable for a long time

<table>
<thead>
<tr>
<th>Date</th>
<th>Sample</th>
<th>Result (g/100g)</th>
<th>Mean:</th>
<th>St Dev:</th>
<th>DF:</th>
</tr>
</thead>
<tbody>
<tr>
<td>10.02.2008</td>
<td>27</td>
<td>22.5</td>
<td>23.1</td>
<td>0.06</td>
<td>13</td>
</tr>
<tr>
<td>16.02.2008</td>
<td>27</td>
<td>21.8</td>
<td>23.1</td>
<td>0.06</td>
<td>13</td>
</tr>
<tr>
<td>26.02.2008</td>
<td>27</td>
<td>22.4</td>
<td>23.1</td>
<td>0.06</td>
<td>13</td>
</tr>
<tr>
<td>7.03.2008</td>
<td>27</td>
<td>23.6</td>
<td>23.1</td>
<td>0.06</td>
<td>13</td>
</tr>
<tr>
<td>17.03.2008</td>
<td>27</td>
<td>23.9</td>
<td>23.1</td>
<td>0.06</td>
<td>13</td>
</tr>
<tr>
<td>27.03.2008</td>
<td>27</td>
<td>23.4</td>
<td>23.1</td>
<td>0.06</td>
<td>13</td>
</tr>
<tr>
<td>6.04.2008</td>
<td>27</td>
<td>23.7</td>
<td>23.1</td>
<td>0.06</td>
<td>13</td>
</tr>
<tr>
<td>16.04.2008</td>
<td>27</td>
<td>23.9</td>
<td>23.1</td>
<td>0.06</td>
<td>13</td>
</tr>
<tr>
<td>26.04.2008</td>
<td>27</td>
<td>22.1</td>
<td>23.1</td>
<td>0.06</td>
<td>13</td>
</tr>
<tr>
<td>6.05.2008</td>
<td>27</td>
<td>22.1</td>
<td>23.1</td>
<td>0.06</td>
<td>13</td>
</tr>
<tr>
<td>16.05.2008</td>
<td>27</td>
<td>23.2</td>
<td>23.1</td>
<td>0.06</td>
<td>13</td>
</tr>
<tr>
<td>26.05.2008</td>
<td>27</td>
<td>22.2</td>
<td>23.1</td>
<td>0.06</td>
<td>13</td>
</tr>
</tbody>
</table>

Within-lab reproducibility $s_{RW}$
Pooled Standard Deviation

- General formula:

\[ s_{\text{pooled}} = \sqrt{(n_1 - 1)s_1^2 + (n_2 - 1)s_2^2 + \ldots + (n_k - 1)s_k^2 \over n_1 + n_2 + \ldots + n_k - k} \]

- **Symbols:**
  - \( k \) number of groups (in this case samples)
  - \( s_1, s_2, \ldots \) within group standard deviations
  - \( n_1, n_2, \ldots \) numbers of measurements made with different samples

---

### Determining precision when sample is not stable for a long time

**Pooled Std Dev**

Determination of protein content

<table>
<thead>
<tr>
<th>Date</th>
<th>Sample</th>
<th>Result (g/100g)</th>
<th>Sample</th>
<th>St dev g/100g</th>
<th>Comp.</th>
</tr>
</thead>
<tbody>
<tr>
<td>10.02.2008</td>
<td>1</td>
<td>10.2</td>
<td>1</td>
<td>0.354</td>
<td>0.125</td>
</tr>
<tr>
<td>10.02.2008</td>
<td>2</td>
<td>13.4</td>
<td>2</td>
<td>0.572</td>
<td>0.980</td>
</tr>
<tr>
<td>10.02.2008</td>
<td>3</td>
<td>17.6</td>
<td>3</td>
<td>0.473</td>
<td>0.447</td>
</tr>
<tr>
<td>13.02.2008</td>
<td>1</td>
<td>10.7</td>
<td>4</td>
<td>0.681</td>
<td>0.927</td>
</tr>
<tr>
<td>13.02.2008</td>
<td>2</td>
<td>14.2</td>
<td>5</td>
<td>0.707</td>
<td>0.500</td>
</tr>
<tr>
<td>13.02.2008</td>
<td>4</td>
<td>18.9</td>
<td>6</td>
<td>0.663</td>
<td>1.320</td>
</tr>
<tr>
<td>18.02.2008</td>
<td>2</td>
<td>12.9</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>18.02.2008</td>
<td>3</td>
<td>16.7</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>18.02.2008</td>
<td>6</td>
<td>12.1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>25.02.2008</td>
<td>6</td>
<td>13.5</td>
<td>1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>25.02.2008</td>
<td>4</td>
<td>17.2</td>
<td>2</td>
<td></td>
<td></td>
</tr>
<tr>
<td>25.02.2008</td>
<td>5</td>
<td>19.2</td>
<td>3</td>
<td></td>
<td></td>
</tr>
<tr>
<td>4.03.2008</td>
<td>2</td>
<td>13.1</td>
<td>4</td>
<td></td>
<td></td>
</tr>
<tr>
<td>4.03.2008</td>
<td>6</td>
<td>12.9</td>
<td>5</td>
<td></td>
<td></td>
</tr>
<tr>
<td>8.03.2008</td>
<td>3</td>
<td>17.4</td>
<td>6</td>
<td></td>
<td></td>
</tr>
<tr>
<td>8.03.2008</td>
<td>4</td>
<td>18.2</td>
<td>7</td>
<td></td>
<td></td>
</tr>
<tr>
<td>8.03.2008</td>
<td>5</td>
<td>13.2</td>
<td>8</td>
<td></td>
<td></td>
</tr>
<tr>
<td>8.03.2008</td>
<td>6</td>
<td>13.5</td>
<td>9</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

\[ s_{\text{pooled}} = 0.598 \text{ g/100g} \]

**Within-lab reproducibility** \( s_{\text{RW}} \)

Different sample matrixes!

But less “long-term”
**u(bias)**

- The *possible bias* of lab’s results from the best estimate of true value is taken into account.

- **u(bias)** can be found:
  - From *repeated* analysis of the same samples with a reference procedure.
  - From *repeated* analysis of certified reference materials (CRMs).
  - From *repeated* interlaboratory comparison measurements.
  - From *repeated* spiking experiments.

Ideally: several reference materials, several PTs because the bias will in most cases vary with matrix and concentration range.

\[
\text{bias}_i = C_{lab_i} - C_{\text{ref}}
\]

\[
RMS_{bias} = \sqrt{\frac{\sum (bias_i)^2}{n}} \quad u(\text{Cref}) = \sqrt{\frac{\sum u(\text{Cref}_i)^2}{n}}
\]

\[
u(bias) = \sqrt{RMS_{bias}^2 + u(\text{Cref})^2}
\]

This component accounts for the average bias of the laboratory results from the reference values.

This component accounts for the average uncertainty of the reference values.
How to conduct a spiking experiment?

• Two analysts determined meropenem (an antibiotic) in blood plasma. Both needed to determine the bias of the procedure. They obtained blank plasma samples and did the following:

• **Analyst 1** took 500 μl of the blank plasma and added 400 μl of methanol. He separated the precipitated proteins by centrifugation and transferred the supernatant into an HPLC vial. He then added 100 μl of meropenem standard solution with suitable concentration to the supernatant and injected the resulting solution into the HPLC system for analysis.

• **Analyst 2** took 500 μl of the blank plasma and added 100 μl of meropenem standard solution. She then added 500 μl of methanol. She separated the precipitated proteins by centrifugation and injected the resulting supernatant into the HPLC system for analysis.

Which analyst did it correctly? Why?

Trueness, bias

Roadmap:

**Possible bias**

\[ u(C_{ref_i}) \text{ from certificates} \]

\[ u(C_{ref_i}) = \frac{s_i}{\sqrt{n_i}} \rightarrow u(C_{ref}) = \sqrt{\frac{\sum u(C_{ref_i})^2}{n}} \]

\[ \text{bias}_i = Clab_i - C_{ref_i} \]

\[ \text{RMS}_{bias} = \sqrt{\frac{\sum (\text{bias}_i)^2}{n}} \rightarrow u(\text{bias}) = \sqrt{\text{RMS}_{bias}^2 + u(C_{ref})^2} \]

**Uncertainty due to random effects**

\[ u(R_w) = s_{RW} \rightarrow u_c = \sqrt{u(R_w)^2 + u(\text{bias})^2} \]
Absolute vs relative uncertainties: Rules of Thumb

- At low concentrations (near detection limit, trace level) use absolute uncertainties
  - Uncertainty is not much dependent on analyte level

- At medium and higher concentrations use relative uncertainties
  - Uncertainty is roughly proportional to analyte level

- In general: whichever is more constant

Available from: http://www.eurachem.org/

Nordtest approach in practice:

Determination of acrylamide in snacks by LC-MS

- Concentration level 998 μg/kg

- Laboratory has analysed two certified reference materials (CRMs) with similar matrixes
  - Potato chips and crisp bread
  - The crisp bread CRM is also used as a control sample
Certified reference material (CRM)

- The **crisp bread** CRM has the following acrylamide content:

  \[ C_{\text{acrylamide}} = (1179 \pm 68) \, \mu g/kg \quad (k = 2, \text{ norm.}) \]

- The **potato chips** CRM has the following acrylamide content:

  \[ C_{\text{acrylamide}} = (860 \pm 42) \, \mu g/kg \quad (k = 2, \text{ norm.}) \]

\[ \text{Days} \quad C \,(\mu g/l) \]

<table>
<thead>
<tr>
<th>Crisp bread</th>
<th>Potato chips</th>
</tr>
</thead>
<tbody>
<tr>
<td>Days</td>
<td>C (mg/l)</td>
</tr>
<tr>
<td>5.01.2008</td>
<td>1172</td>
</tr>
<tr>
<td>6.03.2008</td>
<td>1186</td>
</tr>
<tr>
<td>3.04.2008</td>
<td>1153</td>
</tr>
<tr>
<td>8.01.2009</td>
<td>1151</td>
</tr>
<tr>
<td>18.03.2009</td>
<td>1181</td>
</tr>
<tr>
<td>3.04.2009</td>
<td>1147</td>
</tr>
<tr>
<td>11.04.2009</td>
<td>1097</td>
</tr>
<tr>
<td>16.04.2009</td>
<td>1102</td>
</tr>
<tr>
<td>25.04.2009</td>
<td>1162</td>
</tr>
<tr>
<td>3.08.2009</td>
<td>1138</td>
</tr>
<tr>
<td>28.08.2009</td>
<td>1122</td>
</tr>
<tr>
<td>27.11.2009</td>
<td>1191</td>
</tr>
</tbody>
</table>

Mean: \quad 1150 \, \mu g/kg

Std Dev: \quad 31 \, \mu g/kg
Roadmap:

Possible bias
\[ u(C_{ref_i}) \text{ from certificates} \]
\[ u(C_{ref_i}) = \frac{s_i}{\sqrt{n_i}} \]
\[ u(C_{ref}) = \sqrt{\frac{\sum u(C_{ref_i})^2}{n}} \]

bias\(_i \) = Clab\(_i \) − Cref\(_i \)

\[ RMS_{bias} = \sqrt{\frac{\sum (bias_i)^2}{n}} \]
\[ u(bias) = \sqrt{RMS_{bias}^2 + u(C_{ref})^2} \]

Uncertainty due to random effects
\[ u(R_w) = s_{RW} \]

Combined standard uncertainty
\[ u_c = \sqrt{u(R_w)^2 + u(bias)^2} \]

Finding \( u(R_w) \)

\[ u(R_w) = s_{RW} = 31 \, \mu g/kg \]

\[ u(R_w)_{rel} = s_{RW_{rel}} = \frac{31}{1150} \cdot 100 = 2.70 \% \]
Finding $u(bias)$

<table>
<thead>
<tr>
<th>Ref value</th>
<th>$U$ $(k=2)$</th>
<th>$u_c$</th>
<th>Lab result</th>
<th>$u_{c_rel}$</th>
<th>$bias_i$</th>
<th>$bias_{rel}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>µg/kg</td>
<td>µg/kg</td>
<td>µg/kg</td>
<td>µg/kg</td>
<td>%</td>
<td>µg/kg</td>
<td>%</td>
</tr>
<tr>
<td>1179</td>
<td>68</td>
<td>34</td>
<td>1150</td>
<td>2.88</td>
<td>-29</td>
<td>-2.45</td>
</tr>
<tr>
<td>860</td>
<td>42</td>
<td>21</td>
<td>832</td>
<td>2.44</td>
<td>-28</td>
<td>-3.24</td>
</tr>
</tbody>
</table>

$u(C_{ref\_rel}) = 2.67 \%$

$RMS_{bias\_rel} = 2.87 \%$

$u(bias)_{rel} = 3.92 \%$

24.10.2013

Result:

$u_{c\_rel} = 4.8 \%$

$u_c = 48 \mu g/kg$

$U_{rel \ (k=2)} = 9.5 \%$

$U \ (k=2) = 95 \mu g/kg$

- Acrylamide content in the sample

$C_{acrylamide} = (998 \pm 95) \mu g/kg \ (k = 2, \text{ norm.})$
Choosing the approach

• If you have
  – Competence and time
  – Data on all important influencing quantities
    • Use the Modeling approach

• If you have
  – Quality control data and results of participation in ILC-s or CRM analysis
    • Use the Single-lab validation approach

• Interlab approaches are not generally recommended


Credits

• Some slides from this presentation have been created in collaboration with Bertil Magnusson (SP, Sweden) and used in the training materials.


Thank you for your attention!

• The Online Course of Measurement Uncertainty Estimation in Analytical Chemistry is available from: sisu.ut.ee/measurement/

• You are always welcome to contact me: ivo.leito@ut.ee