Uncertainty arising from sampling

S Ellison (LGC Limited)
M Ramsey (University of Sussex)

Introduction

• Does measurement uncertainty include sampling?
• How to estimate uncertainties from sampling
• Uncertainties from sampling in the food sector
Measurement uncertainty

ISO definition
“A parameter, associated with the result of a measurement, that characterises the dispersion of the values that could reasonably be attributed to the measurand”

22.7 ± 4.8 g

The part of the result after the ±

Does measurement uncertainty include sampling?

EURACHEM position
• If the measurand relates to a bulk material from which samples are taken for analysis, the uncertainty in the estimated value for the measurand must include the uncertainty arising from the sampling process

• If the result is reported on the sample ‘as received’ by the laboratory, only within-laboratory sub-sampling contributes to the uncertainty
Examples

- Measuring and reporting the amount of pesticide in a laboratory sample of capsicum (green peppers)
  - Little or no sampling/subsampling
  - Sampling is not part of the measurement process

- Reporting the average level of pesticide in the bulk container (consignment) from which the laboratory sample was taken
  - Sampling may greatly affect the reported result
  - Sampling uncertainty matters

Estimating sampling uncertainty
Different approaches to control of sampling

• Gy: Well respected, based on management and control to eliminate sampling uncertainties

• Sampling uncertainties quantified using replication
  – Ramsey et al
  – Eurachem Guide

• Applying modelling approaches to sampling uncertainty
  – Minkkinen et al

<table>
<thead>
<tr>
<th>Method #</th>
<th>Method description</th>
<th>Sampler s (People)</th>
<th>Protocol s</th>
<th>Component estimated</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Duplicates</td>
<td>single</td>
<td>single</td>
<td>Yes No Yes No 1</td>
</tr>
<tr>
<td>2</td>
<td>Multiple protocols</td>
<td>single</td>
<td>multiple</td>
<td>between protocols Yes No 1</td>
</tr>
<tr>
<td>3</td>
<td>CTS</td>
<td>multiple</td>
<td>single</td>
<td>between samplers Yes Yes 2</td>
</tr>
<tr>
<td>4</td>
<td>SPT</td>
<td>multiple</td>
<td>multiple</td>
<td>between protocols +between samplers Yes Yes 2</td>
</tr>
</tbody>
</table>

CTS = Collaborative Trial in Sampling, and SPT = Sampling Proficiency Test.

Simplest Empirical method is ‘Duplicate Method’
Using the ‘duplicate method’
1) Separating sampling and analysis

Sampling target:
Portion of material, at a particular time, that the sample is intended to represent.

Using the ‘duplicate method’
2) Replicating sampling

.... to at least 8 sampling targets
Statistical Analysis for the duplicate method

- Fully nested, balanced 2-way layout
- Analysis of variance gives sampling and analytical variance
  - each corresponding to a relevant standard uncertainty contribution
- Robust analysis of variance (RANOVA) suggested for outlier-contaminated data

Example: Nitrate in lettuce (Eurachem Guide p 35ff)

20,000 lettuce heads

1 “sampling target”

每一个湾被采样

- Every bay sampled
- Decision for each bay
Example 1: Duplicate method

8 (or more) targets sampled in duplicate

Duplicate sampling arrangement

Example: Analysis

8 sampling targets
Sampled in duplicate
Each sample duplicate analysed in duplicate
Example: Results

<table>
<thead>
<tr>
<th>SAMPLE</th>
<th>S1A1</th>
<th>Uncertainty</th>
<th>x - U</th>
<th>x + U</th>
<th>Probabilistic Classification</th>
</tr>
</thead>
<tbody>
<tr>
<td>TARGET</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>A</td>
<td>3898</td>
<td>639</td>
<td>3259</td>
<td>4537</td>
<td>Poss Cont</td>
</tr>
<tr>
<td>B</td>
<td>3910</td>
<td>641</td>
<td>3269</td>
<td>4551</td>
<td>Poss Cont</td>
</tr>
<tr>
<td>C</td>
<td>5708</td>
<td>936</td>
<td>4772</td>
<td>6644</td>
<td>Cont</td>
</tr>
<tr>
<td>D</td>
<td>5028</td>
<td>825</td>
<td>4203</td>
<td>5853</td>
<td>Prob Cont</td>
</tr>
<tr>
<td>E</td>
<td>4640</td>
<td>761</td>
<td>3879</td>
<td>5401</td>
<td>Prob Cont</td>
</tr>
<tr>
<td>F</td>
<td>5182</td>
<td>850</td>
<td>4332</td>
<td>6032</td>
<td>Prob Cont</td>
</tr>
<tr>
<td>G</td>
<td>3028</td>
<td>497</td>
<td>2531</td>
<td>3525</td>
<td>Uncont.</td>
</tr>
<tr>
<td>H</td>
<td>3966</td>
<td>650</td>
<td>3316</td>
<td>4616</td>
<td>Poss Cont</td>
</tr>
</tbody>
</table>

Example 1: Results

Robust ANOVA:
- $s_{samp} = 319.05 \text{ mg kg}^{-1}$
- $s_{anal} = 167.94 \text{ mg kg}^{-1}$
- $s_{meas} = \sqrt{(s_{samp}^2 + s_{anal}^2)} = 360.55 \text{ mg kg}^{-1}$

Classical ANOVA:
- $s_{samp} = 518.2$, $s_{anal} = 148.2$; $s_{meas} = 538.9 \text{ mg kg}^{-1}$
Sampling uncertainties in food analysis

Review of sampling uncertainties in foods: Overview

- Collate available data from literature
  - 23 sources identified
  - 30 product types
  - 59 product/analyte combinations
  - 13 products in retail environments; 17 factory/wholesale
- Apply duplicate method to increase data set on foods of interest
  - A further 16 product/analyte combinations
- Review sampling uncertainties for trends
Results: Sampling/analytical ratios

Most sampling uncertainties are larger than analytical precision

Some evidence of trend with analyte concentration

Trends with concentration
i) Raw results

Sampling RSD increases with decreasing analyte concentration
**Trends with concentration**

i) log-log relationship

A Horwitz-like function describes sampling SD.

The relationship is very approximate.

**Limitations**

- Literature surveys reflect interest in sampling
  – Possibly biased towards known problems

- Experimental work chosen to provide a range of examples
  – not a random sample of sampling

- Values differ from fitted line by about ±1 in log_{10}:
  – Approximately one order of magnitude

- Sampling variation need not follow any particular distribution
Practical implementation

- The duplicate method requires a minimum of 8 replicated sampling targets, or 16 sampling exercises
  - Economical only when many more increments are normally taken and measured separately

- Most useful when developing or comparing proposed sampling strategies in practice?

Conclusions

- Primary sampling from the bulk contributes to the uncertainty when the measurand is defined as a property of the bulk material
- Relatively economical empirical approaches to estimating sampling uncertainty are available
- Sampling uncertainties are often considerably larger than analytical uncertainty
- Available data suggest that sampling standard deviation can be predicted to approximately an order of magnitude
Acknowledgements

- UK Food Standards Agency
  Supported the research reported here

- UK National Measurement office
  International programme