Uncertainty from sampling food—an empirical approach

Michael Thompson
Birkbeck University of London
Malet Street, London WC1E 7HX
m.thompson@bbk.ac.uk

Michael H Ramsey
University of Sussex
Brighton BN1 9QG
m.h.ramsey@sussex.ac.uk

Scientists involved in this project

From the University of Sussex:
Prof Michael Ramsey
Dr Bas Geelhoed
Dr Jennifer Lyn
Ilaria Palestra

From Birkbeck University of London:
Prof Michael Thompson

From the Food Standards Agency:
Dr Roger Wood
Dr Andy Damant
Dr Ellis Evans
‘Quality’ concepts and actions

- **Fitness for purpose**: what uncertainty is best for the customer?
- **Method validation**: can the method produce a suitably low uncertainty?
- **Internal quality control**: have things changed since validation? *(i.e., did the method work well on the day?)*
- **Proficiency testing**: does the whole system really work?

How it all fits together

- **Fitness for purpose**: What accuracy best suits end-user needs?
  - **Expert opinion**
  - **Decision theory**

- **Method validation**: What accuracy can my method supply? Is it fit for purpose?
  - **Collaborative trials**
  - **Precision studies**
  - **Reference targets**

- **Quality control**: Does my method consistently remain fit for purpose?
  - **Duplication**
  - **Control charts**
  - **Control targets**
  - **Proficiency tests**
The traditional approach—sampling considered separately from measurement.

- Design ‘correct’ sampling protocol to give a ‘representative’ sample.
- Train sampler to apply the protocol.
- Assume $u_{sam} = 0$.

The traditional approach is logically untenable. Why?

- Customers (and other stakeholders) need to know the total combined uncertainty to make informed decisions about the target.

$$u = \sqrt{u_{sam}^2 + u_{an}^2}$$
Uncertainty in compliance decisions

Combining uncertainties

\[ u = \sqrt{u_{sam}^2 + u_{an}^2} \]
Components of sampling uncertainty

- **Bias**—difficult, often impracticable to address.
- **Precision**—Easy to address so long as a random element can be introduced into replicating the procedure.

Sampling bias

- Some experts think that sampling bias does not exist.
- Essentially they hold that sampling methods are empirical, *i.e.*, give an unbiased sample by definition.
- That is not generally correct—it is easy to see how sampling bias could arise in practice.
One way of taking a biased sample!

Addressing bias

<table>
<thead>
<tr>
<th>Analytical</th>
<th>Sampling</th>
</tr>
</thead>
<tbody>
<tr>
<td>Reference material</td>
<td>Reference target</td>
</tr>
<tr>
<td></td>
<td>(Severe problems with cost, stability)</td>
</tr>
<tr>
<td>Reference method vs. candidate method (multiple test materials)</td>
<td>Reference method vs. candidate method (multiple test targets)</td>
</tr>
</tbody>
</table>
Sampling precision

- Variations in execution of procedure.
- Variations in composition (heterogeneity) of target.
  - Sampling precision may vary from target to target of the same nominal type.
  - Initial validation of the sampling protocol needs to be supported by ongoing checks (internal quality control).
- Good estimation of precision needs **RANDOM** replication of sampling.

Random duplication—sampling to a pattern

Composite Sample A

Composite Sample B
Stratified random design

Sampling from a conveyor belt
Collaborative trial

• Requires:
  multiple targets, multiple samplers, duplicate samples, duplicate analysis (random repeatability conditions).

• Provides:
  analytical repeatability variance, between-sample (repeatability) variance, between-SAMPLER (reproducibility) variance.

• Drawbacks:
  VERY expensive.

• Current usage:
  research only.

Method validation—nitrate in lettuce

• Nitrate a potential risk to human health
• EU threshold is 4500 mg kg\(^{-1}\) for batch concentration
• Current sampling protocol specifies taking 10 heads to make a single composite sample from each batch.
• Sampling uncertainty unknown
Sampling of lettuce from one bay of greenhouse

Randomisation is important…

…but not always exactly feasible. Here we use systematic replication.
Balanced sampling design

Target 1

Result
Result

Target 1

Result
Result

Target 8

Result
Result

Results of the balanced experiment
Statistics from robust ANOVA

\[ \hat{\sigma}_{\text{anal}} = 168 \]
\[ \hat{\sigma}_{\text{samp}} = 319 \]
\[ \hat{\sigma}_{\text{comb}} = 361 \]
\[ \overline{\mu} = 4346 \]

• Is the accuracy fit for purpose?
  • (Note: \( \sigma \) is equivalent to standard uncertainty if measurement bias is absent.)

Fitness for purpose

• A result is fit for purpose when it maximises its expected utility.
• This means roughly that we need to minimise expected costs in the long term.
• There are operational costs of sampling and analysis.
• There are potential costs resulting from incorrect decisions based on the result.
• Both of these costs depend on uncertainty.
The cost of accuracy, $L_m = f(u)$

Cost of incorrect decisions ($L_d$):
1—probability of false rejection
Probabilities of false acceptance

Typical loss function

Average loss = Cost of incident \times \text{probability of incident}
Long-term loss

Fit-for-purpose uncertainty
Total cost $T$

\[ T = L_m + L_d \]

\[ = L_m + \int \int L(x, \mu) P_m(x|\mu) P(\mu) \, dx \, d\mu \]

$P_m(x|\mu)$ is the distribution of the result, given the true value;

$P(\mu)$ is the distribution of knowledge about where the true value might lie.

Loss function for sampling lettuce
Getting near-optimal uncertainty

\[
\begin{align*}
\hat{\sigma}_{\text{anal}} &= 168 \\
\hat{\sigma}_{\text{samp}} &= 319 \\
\hat{\sigma}_{\text{comb}} &= 361 \\
\end{align*}
\]

\[
\begin{align*}
\sigma_{\text{anal}} &= 168 \\
\sigma_{\text{samp}} &= 160 \\
\sigma_{\text{comb}} &= 232 \\
\end{align*}
\]

Original: 10 increments

Proposed: 40 increments

Internal quality control

- Set up control chart for \(d\) (Shewhart or J-chart) using control lines at 0, \(\pm 2s\), and \(\pm 3s\), where
- For routine or occasional use (for sampling and analysis combined).
- **Note:** a result may be unfit for purpose, even if the error is due to heterogeneity and not the method.
Combined sampling/analysis Shewhart control chart—aluminium in animal feed

Proficiency test in Sampling

Assigned value ideally from an more reliable and independent sampling method
Desiderata

- Samplers use their own preferred sampling protocol.
- Scheme provider conducts analysis under repeatability conditions (with $\sigma_a \ll \sigma_{s(R)}$).
- Provider specifies a fitness-for-purpose criterion.
- Provider uses an independent assigned value if possible
- Provider calculates a z-score.

Practical points

- Sampling must be “replicable” and “unobserved”.
- There may be overall target-specific bias if assigned value is a consensus.
- Expensive.
- Usage: research only at present
General references

• Measurement uncertainty arising from sampling: a guide to methods and approaches.
  M H Ramsey and SLR Ellison
  Eurachem/Eurolab/CITAC/Nordtest/ AMC Guide
  2007, 111 pages.
  (Free download from www.eurachem.org/guides/UfS_2007.pdf)

• Uncertainty from sampling, in the context of fitness for purpose. (Review)
  M H Ramsey and M Thompson