Use of error modelling in evaluating measurement uncertainty

Adriaan M.H. van der Veen

Nederlands Meetinstituut

Maurice G. Cox

National Physical Laboratory

“Mainstream GUM” (1)

1. Express measurand Y mathematically in input quantities X_i; function f should contain <u>all</u> variables that could contribute significantly to the measurement uncertainty
2. Determine all x_i
3. Determine all $u(x_i)$
4. Evaluate covariances

“Mainstream GUM” (2)

5. Calculate *Y*
6. Calculate the combined standard uncertainty
7. Select coverage factor *k*
8. Report expanded uncertainty *U*

Step 1 – Modelling

- Purpose: to translate measurement process into mathematical formalisms
- Formalisms required for processing numerical data
- Model should encompass all influencing factors
- Model relates influencing factors to measurand

Modelling process

- Different approaches possible
- Often an augmentation of an existing relationship defining the measurand works fine
- ‘Cascading’ approach (e.g. Eurachem Guide) keeps maths simple, unless … correlations come into play

Error modelling

- Systematic approach to the problem
- In particular useful when starting from scratch
- Technique is very
  - flexible
  - robust
  - forces to state assumptions

Error and uncertainty (I)

- Existence of *errors* reason for having *uncertainty*
- Statistical treatment of errors provides corrections and standard uncertainties
- A single value for an error does not tell much about the uncertainty component …
- … but the behaviour does

Definition

- An error is defined as the departure of a variable from its expectation
- Expectation is defined as the result of

$$\mu = \int_{-\infty}^{\infty} x f(x) dx$$

- Relation with a true value requires an additional assumption!

Error and uncertainty (II)

- Relationship is given through the variance

$$Var(X) = E\left(\left(X - \mu\right)^2\right) = E\left(\varepsilon^2\right)$$

- Pdf describes behaviour of errors

Example

- Given
  - $<q> = 10$
  - $\mathrm{Var}(q) = \sigma^2 = 2$
- Estimates
  - m 10.153
  - s^2 1.49

Drawing (i)	q_i	δq_i
1	8.22	-1.78
2	10.63	0.63
3	10.82	0.82
4	8.71	-1.29
5	10.51	0.51
6	11.27	1.27
7	11.66	1.66
8	9.19	-0.81
9	11.33	1.33
10	9.19	-0.81

What does the GUM tell us?

- Propagation of uncertainties
  - clause 5
  - useful when directly modelling uncertainties
- Propagation of errors
  - clause E.3
  - basis of the GUM
  - conversion of an error model into an uncertainty model

Error modelling

- Each quantity q can be expressed in its expectation and its error, e.g.

$$q = \langle q \rangle + \delta q,$$

- Using the error propagation formula, the uncertainty in q equals the uncertainty of its error

Case: degrees of equivalence

- Given a laboratory comparison operated with a batch of artefacts
- A degree of equivalence is defined as
  - the difference (of a lab's result) from the reference value
  - the uncertainty thereof at 95% level of confidence

Expression for the deviation

- For a laboratory *i*

$$D_i = x_i - x_{ref} \,,$$

- where x_i denotes the laboratory result and x_{ref}.
- Between 2 laboratories *i* and *j*

$$D_{ij} = (x_i - x_{ref}) - (x_j - x_{ref}) = x_i - x_j \,,$$

Modelling

- It seems that both definitions are fine for expressing the uncertainty of D_i and D_{ij} respectively, but what about
  - batch inhomogeneity
  - stability of the artefacts

Batch inhomogeneity

- Affects *D* as a laboratory receives only one artefact
- Can be defined as a correction (to the reference value)

$$x_{bb,i} = \langle x_{bb,i} \rangle + \delta x_{bb,i}.$$

- each artefact has its own error (they are not necessarily equal!)

Stability

- Instability may be due to
  - intrinsic instability
  - storage conditions
  - (less than ideal) transport conditions
- Issues affect each artefact individually

$$x_{stab,i} = \langle x_{stab,i} \rangle + \delta x_{stab,i}.$$

Consensus value

- A consensus value is used as reference

$$x_{cv} = \langle x_{cv} \rangle + \delta x_{cv}.$$

- Including all corrections leads to

$$x_{cv,i} = x_{cv} + x_{bb,i} + x_{stab,i} = \ldots$$

$$\ldots = \langle x_{cv} \rangle + \delta x_{cv} + \langle x_{bb,i} \rangle + \delta x_{bb,i} + \langle x_{stab,i} \rangle + \delta x_{stab,i}.$$

Assumptions

- So far, no assumptions have been made other than that inhomogeneity and instability of artefacts play a role
- Before further developing the model, it is now useful to make some assumptions/assertions

Corrections

- It is assumed that the homogeneity study is implemented in such a way, that the correction is 0.
- It is assumed that the stability study is implemented in such a way, that the correction is 0.

Consensus value

- Consensus value depends on laboratory results

$$x_{cv} = \sum w_i x_i .$$

- In terms of errors,

$$x_{cv} = \sum w_i \langle x_i \rangle + \sum w_i \delta x_i .$$

- where δx_i is the aggregated laboratory error term

Deviation from x_{cv}

- Model can now be expressed as

$$D_i = \langle x_i \rangle + \delta x_i - \left(\langle x_{cv} \rangle + \delta x_{cv} + \delta x_{bb,i} + \delta x_{stab,i} \right).$$

- where it should be noted that δx_{cv} and δx_i are (cor)related!
- Correlation can be dealt with by
  - substitution of expression for δx_{cv}
  - evaluation of the covariance directly

$$u(x_i, x_{cv}) = w_i u^2(x_i),$$

Degree of equivalence D_i

- Expression for uncertainty

$$u^2(D_i) = u^2(x_i) + u^2(x_{cv}) + u_{bb}^2 + u_{stab}^2 - 2u(x_i, x_{cv}),$$

- follows directly from using the error propagation formula (GUM, E.3)

Bilateral case

- Using the same methodology,

$$D_{ij} = \left\{x_i - \left(\langle x_{cv}\rangle + \delta x_{cv} + \delta x_{bb,i} + \delta x_{stab,i}\right)\right\} -$$
$$\left\{x_j - \left(\langle x_{cv}\rangle + \delta x_{cv} + \delta x_{bb,j} + \delta x_{stab,j}\right)\right\} = \ldots$$
$$\ldots = x_i - \left(\delta x_{bb,i} + \delta x_{stab,i}\right) - x_j + \left(\delta x_{bb,j} + \delta x_{stab,j}\right).$$

- which leads to

$$u^2\left(D_{ij}\right) = u^2\left(x_i\right) + u^2\left(x_j\right) + 2u_{bb}^2 + 2u_{stab}^2.$$

Important notes on D_{ij}

- Expectations of consensus value vanish
- Error terms of consensus value cancel
- Between-bottle variation enters *twice*
- Instability also enters *twice*

Elaboration on x_{cv}

- Error term is aggregated (components: δx_i)
- Error terms δx_i also aggregated
- Aggregation irrelevant so far
- However: when combining results from different comparisons, aggregated terms become crucial

Aggregation and de-aggregation

- Aggregation
  - Useful in many practical applications
  - GUM allows aggregated terms in the model
- De-aggregation
  - Necessary to address covariance-issues
  - Process answers the question: “What is contained in it?”
  - Knowledge counts, not the magnitude of the uncertainties!

De-aggregating δx_{cv}

- Separation in a within- and between laboratory term

$$\delta x_i = \delta x_{(L)i} + \frac{1}{n_i} \sum \delta x_{(r)i},$$

- Between-laboratory term still aggregated
- Uncertainty components become inseparable when $n_i = 1$

Expression for δx_{cv}

- De-aggregation leads to

$$\delta x_{cv} = \sum w_i \delta x_{(L)i} + \sum w_i \left(\frac{\sum \delta x_{(r)i}}{n_i} \right).$$

- and finally to

$$u^2\left(x_{cv}\right) = \sum w_i^2 u^2\left(x_{(L)i}\right) + \sum w_i^2 \left(\frac{\sum u^2\left(x_{(r)i}\right)}{n_i^2} \right),$$

Relationship with ISO 5725

- Assuming $n_i = n$,
- each $u(\delta x_{(L)}) = s_l$,
- each $u(\delta x_{(L)}) = s_r$,

$$u^2(x_{cv}) = s_L^2 \sum w_i^2 + \left(\frac{s_r^2}{n} \right) \sum w_i^2 .$$

- and if all weights are equal (= 1/p), the one-way ANOVA expression from ISO 5725-2 is obtained

“Top-down” vs. “bottom-up”

- Distinction is not very useful at technical level
- Approaches should be consistent
- Aggregated terms have great benefits, but are difficult to handle if it is unknown what is contained
- De-aggregation is key to demonstrating compliance with ISO/IEC 17025

Concluding remarks (1)

- Error modelling is a useful technique for constructing uncertainty models
- Main benefits are
  - ease of use
  - clear understanding of the assumptions made
  - often shortest route to the desired uncertainty model

Concluding remarks (2)

- Error modelling in particular useful when
  - starting from scratch
  - augmenting existing models
- Mechanism of conversion is by means of error propagation (GUM, E.3), rather than uncertainty propagation