

TOWARDS A REVISED GUIDE TO THE EXPRESSION OF UNCERTAINTY IN MEASUREMENT

Maurice Cox
National Physical Laboratory, UK
maurice.cox@npl.co.uk

AMCTM, Paris, 23–25 June 2008

Outline

Working Group 1 (WG1), **Expression of uncertainty in measurement**, of Joint Committee for Guides in Metrology (JCGM)

Steps being taken towards a revised GUM

Contents and strategy for change

Reasons for revision of current GUM

Mix of Frequentist and Bayesian treatments

Internal inconsistencies, and external inconsistencies with GUM Supplement 1, other JCGM/WG1 documents, and the VIM

Inadequacies: present GUM does not cover models

- where the input quantities have related finite degrees of freedom **and** have dependencies
- with several output quantities
- with complex quantities

Development

Align the GUM with today's thinking on uncertainty evaluation

Resist change that risked alienating the user community

Provide **practical prescriptions** for uncertainty evaluation for a wider set of problems **and** having a sound theoretical basis

Identify class of problems that would be handled **reliably** by the (existing) law of propagation of uncertainty

Measurement model and probability distributions

Revised version would also be based on measurement model relating **all** quantities involved in a measurement:

- Measurable quantities
- Correction effects
- Etc.

All quantities characterized by **probability distributions**

Reliability

Provision of **reliable** uncertainties associated with estimates of quantities given by model not necessarily a simple task

Reliability: closeness of computed values to actual solution **given** model and distributions for input quantities

One significant decimal digit (two in demanding situations) considered reliable in terms of **fitness for purpose**

To assure one significant decimal digit might require sophisticated calculation but, for a result obtained by a simple means, the numerical accuracy might be **unknown**

No different from many other scientific situations in terms of the **balance** of simplicity and sophistication

Modelling the measurement

Attention would be given to the types of model that often arise:

1. Models with **one or several measurands** (output quantities)
2. **Explicit** models or **implicit** models
3. Models involving **complex quantities**

Above classification valuable in formulating and solving uncertainty evaluation problems [[Cox & Harris, 2006](#)]

Probabilistic considerations

Bayesian approach [Lindley, 2006], accounting for relevant work [Gleser, 1998; Kacker and Jones, 2003; Weise and Wöger, 1992]

Advantage: additional knowledge of quantities in model, e.g. that measurand is positive, can be **incorporated** [Elster, 2007]

Provide **prescriptions** for obtaining a probability distribution for output quantities knowing model and distributions for input quantities (and possibly prior distributions for output quantities)

Advantages

Each quantity in model regarded as **random variable** characterized by probability distribution

Distribution **encodes knowledge** available about quantity

Consistent with current GUM, but distinction between Type A and B evaluations of uncertainty **no longer necessary**

Apply **Bayes' theorem** [Gelman et al, 2004] or **Principle of Maximum Entropy** [Weise and Wöger, 1992] to obtain probability distribution for an input quantity

Formulæ and prescriptions

Generalization of GUM uncertainty framework to these model types: formulæ to be provided to deliver

- **Estimates** of the output quantities
- **Associated standard uncertainties** or an **uncertainty matrix**

Coverage intervals or regions

Procedures to determine, for prescribed coverage probability p , **coverage interval or region** for output quantities

Rely on GUM **assumption** that the output quantities can be characterized by a (joint) Gaussian or t -distribution

Coverage interval from expanded uncertainty

Vector measurand: extension to coverage **regions**

Bottom-up and top-down models

Form of model used in existing GUM considered a **bottom-up** model: the model relates to a set of specified input quantities

Consideration in revised GUM to **top-down** models: link to work of ISO Committee TC/69, **Application of statistical methods**

TC/69 Standard **ISO 21748:2004** concentrates on top-down models: a model of the measurement **not generally available**

Set of values, reasonably attributed to measurand, described by values for **repeatability, reproducibility and trueness** of method in use, obtained by collaborative studies [**ISO 5725-2:1994**]

Bottom-up and top-down models (continued)

These values correspond to those of the **intra- and inter-laboratory components** of standard uncertainties, together with standard uncertainty associated with **trueness** of method

Standard uncertainty associated with a measurand estimate evaluated by **combining** these standard uncertainties as prescribed by existing GUM

This (combined) standard uncertainty **augmented** by standard uncertainties related to any influences not covered by collaborative study, and so regarded as an outcome of a composite top-down and bottom-up model

Implementations of propagation of distributions

Uncertainty evaluation basis: **propagation of distributions**

Propagates probability distributions for input quantities through model to furnish **distributions for output quantities**

All uncertainty evaluation methods, including GUM uncertainty framework, are **approximations** to propagation of distributions

Implementations of the propagation of distributions (continued)

Revised GUM to contain or refer to prescriptions for particular implementations:

- GUM uncertainty framework
- MCM
- Monte Carlo Markov chains (MCMC) [[Gelman et al, 2004](#)]

When used **appropriately**, MCM and MCMC can deliver solutions with **controlled reliability**

MCM and MCMC

MCM approximates distribution for output quantities by evaluating model for **draws from distributions for input quantities**

MCMC constructs a (Markov) **chain of random variables** from some starting point

When members in sequence **no longer** depend on starting point, they represent joint distribution for output quantities

For MCM and MCMC, coverage intervals or regions determined **without** making the GUM normality assumption

Conclusions

Transition from the existing GUM to revised GUM considered

Modelling and probabilistic aspects of revised GUM emphasized

Much of the material based on the work of JCGM/WG1, but the author takes responsibility for the content