The Joint Committee for Guides in Metrology, WG1: trying to establish some certainty in measurement uncertainty

Walter Bich
INRIM and JCGM-WG1

Disclaimer

The opinions expressed in the following are personal views. They are not necessarily shared by other JCGM-WG1 members, and certainly do not represent a position of JCGM-WG1. Some notable exceptions will be highlighted

Framework

- 1977-79 BIPM questionnaire on uncertainties
- 1980 Recommendation INC-1
- 1981 Establishment of WG3 on uncertainties under ISO TAG4: BIPM, IEC, IFCC, ISO, IUPAC, IUPAP, OIML
- 1981 Recommendation CI-1981
- 1986 Recommendation CI-1986
- 1993 Guide to the expression of uncertainty in measurement GUM
- 1995 Reprint with minor corrections
- 1997 Establishment of the Joint Committee for Guides in Metrology JCGM. ILAC joins in 1998

Joint Committee for Guides in Metrology

- Present Chair: the BIPM Director

- The JCGM has two working groups (WGs)

• WG 1 has responsibility for the Guide to the expression of uncertainty in measurement, GUM

• WG 2 has responsibility for the International vocabulary of metrology, VIM

WG1 published documents

JCGM 101:2008 (Monte Carlo)

JCGM 100:2008 (GUM 1995 with minor corrections)

JCGM 104:2009 (Introduction to uncertainty)

JCGM 102:2011 (Any number of output quantities)

JCGM 106:2012 (Conformity assessment)

JCGM GUM-6:2020 (Modelling)

All these are published by BIPM, OIML, ISO and IEC. Adopted by IFCC, ILAC, IUPAC and IUPAP

WG1 forthcoming documents

JCGM GUM-1 (Introduction). Will replace JCGM 104:2009

JCGM GUM-5 (Examples)

These two, as well as JCGM GUM-6:2020, are published under a new perspective, in which “the GUM” is the whole suite of documents

WG1 and WG2

The two WGs pursue the goal of maximum harmonisation and consistency

Measurement uncertainty is a topic of common interest

WG1 is working on a set of definitions related to MU

There are some first results

Measurement uncertainty

doubt about the true value of the measurand that remains after making a measurement*

*This *is* a JCGM-WG1 position

MU as *doubt*

MU is a (subjective) state of mind

doubt about the true value of the measurand that remains after making a measurement

MU is *the* concept, not a measure of the concept

MU as *doubt**

Doubt is not in nature, it is in the experimenter's mind. Of course, it is solidly, but not exclusively based on objective data. The desire to remove subjectivity from measurement, and from science at large, is just a hope. Complete objectivity in science is gibberish

MU is the concept, not a measure of the concept**. A nice decoupling between the concept and its measures

This decoupling allows the definition to be easily transferred to a corresponding definition for *examination uncertainty* (for nominal properties)

*Long discussions within JCGM-WG1 on doubt *vs* uncertainty

**Long discussions as well on the convenience of such a change with respect to the past

Measurement uncertainty

doubt about the **true value** of the measurand that remains after making a measurement

The true value, for a long time demonised as an idealised concept, is a key element in estimation theory as based on probability and statistics

The true value is strictly connected with the concept of real number

Philosophy of science cannot prescind from or come before mathematics

Measurement uncertainty

doubt about **the** true value of the measurand that remains after making a measurement*

The, not ***a*** true value. Uniqueness!!! There is not a whole set of true values (consistent with the definition of a quantity)

GUM 1995 ambiguities

True value:
D.3.4 … Thus in this case [the thickness of a sheet], because of an incomplete definition of the measurand, the “true” value has an uncertainty that can be evaluated from measurements made at different places on the sheet. At some level, every measurand has such an “intrinsic” uncertainty that can in principle be estimated in some way.

This leads to the VIM3 definition:
quantity value consistent with the definition of a quantity

In turn leading to:
there is not a single true quantity value but rather a set of true quantity values consistent with the definition

…and to definitional uncertainty

But…

1.2 This Guide is primarily concerned with the expression of uncertainty in the measurement of a well-defined physical quantity — the measurand — that can be characterized by an essentially unique value.

If the phenomenon of interest can be represented only as a distribution of values or is dependent on one or more parameters, such as time, then the measurands required for its description are the set of quantities describing that distribution or that dependence.

D.3.5 The term "true value of a measurand" …(often truncated to "true value") is avoided in this Guide because the word "true" is viewed as redundant.

The GUM 1995 does indeed use the true value, although ambiguously

Measurement uncertainty

doubt

about the true value of the measurand

that remains after making a measurement

The doubt is about the measurand (given its estimate), *not* about the estimate

There is a deep divergence in this between the frequentist and the subjective views of probability

Measurement uncertainty

doubt about the true value of the measurand that remains after making a

measurement

«measurement» has a meaning broader than that intended in the VIM

Measurement uncertainty

NOTE 1 Measurement uncertainty can be described fully and quantitatively by a probability distribution on the set of possible values of the measurand. It can be described summarily and approximately by a quantitative indication of the dispersion (or scatter) of such distribution.

Measurement uncertainty

NOTE 2 For scalar measurands, measurement uncertainty can be described summarily by, for example, the *standard uncertainty*, a coverage interval with specified coverage probability, or by selected quantiles of the probability distribution in Note 1. For multivariate measurands, measurement uncertainty can be described, for example, by the covariance matrix or by a coverage region, with specified coverage probability.

Measurement uncertainty

NOTE 3 When a quantitative expression is impractical, measurement uncertainty can be expressed using an ordinal scale of levels of confidence in the assigned value.

Measures of MU

Standard measurement uncertainty

Measurement uncertainty expressed as a standard deviation*

Notes to be decided

A pillar (*the* pillar?) of the INC-1 Recommendation

Personal note: I am happy with the definition,
I only miss the specification of the relevant
random variable or probability distribution

*VIM3 definition

Measures of MU

Standard measurement uncertainty

- The sample standard deviation in Type A evaluations is biased down
- The standard deviation of a random variable might not exist, or be infinite
- It might be little representative

Yet…

Standard measurement uncertainty

Is deeply rooted in current practice

Is the natural companion of the mean

Is the natural consequence of the choice of a specific loss function, the Mean Squared Error

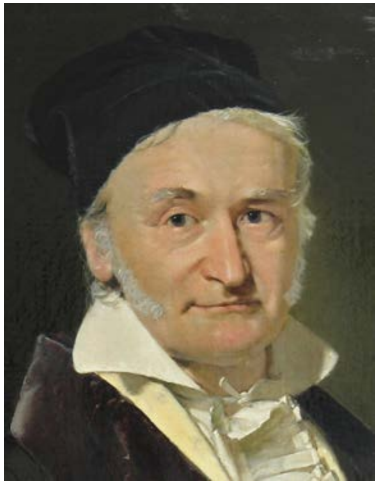
$$\mathrm{MSE}(\hat{\theta}) = \mathrm{E}\left[(\hat{\theta} - \theta)^2\right]$$

or

$$\mathrm{MSE}(\hat{\theta}) = \mathrm{VAR}(\hat{\theta}) + \left[\mathrm{E}(\hat{\theta} - \theta)\right]^2$$

The great man

It is by no means self-evident how much loss should be assigned to a given observation error. On the contrary, the matter depends in some part on our own judgment. Clearly we cannot set the loss equal to the error itself; for if positive errors were taken as losses, negative errors would have to represent gains. The size of the loss is better represented by a function that is naturally positive. Since the number of such functions is infinite, it would seem that we should choose *the simplest function* having this property. That function is unarguably the square, and the principle proposed above results from its adoption.



C. F. Gauss, Theoria Combinationis Observationum Erroribus Minimis Obnoxiae, 1820

… and more…

LAPLACE has also considered the problem in a similar manner, but he adopted *the absolute value of the error* as his measure of loss. Now if I am not mistaken, this convention is no less arbitrary than mine. Should an error of double size be considered as tolerable as a single error twice repeated or worse? Is it better to assign only twice as much influence to a double error or more? The answers are not self-evident, and the problem cannot be resolved by mathematical proofs, but only by an arbitrary decision. Moreover, it cannot be denied that *LAPLACE'S convention violates continuity and hence resists analytic treatment*, while the results that my convention leads to are distinguished by their wonderful simplicity and generality.

C. F. Gauss, Theoria Combinationis Observationum Erroribus Minimis Obnoxiae, 1820

The results that my convention leads to are distinguished by their wonderful simplicity and generality

- Variances propagate through a measurement model in a comparatively simple way
- The rule is distribution-free
- The propagation is exact for linear models

In these unique features lies the reason for the enduring success of Gauss' idea

Measures of MU

expanded measurement uncertainty

product of a combined standard measurement uncertainty and a factor larger than the number one *

Simply a multiple of SU. No added value

*VIM3 definition

Measures of MU

expanded measurement uncertainty

quantity defining an interval about the result of a measurement that may be expected to encompass a large fraction of the distribution of values that could reasonably be attributed to the measurand**

Opens a door to coverage intervals with a stipulated coverage probability

**JCGM 100:2008 definition

A delicate issue. Certainly not a measurable, physical quantity in the sense of the ISQ, International System of Quantities. Uncertainty is not a measurand, and has no uncertainty. Dark uncertainty is a misleading term for an unexplained dispersion.

Straying on forbidden grass…

- As said, the standard uncertainty is the natural companion of the mean
- The mean could not be representative
- Laplace's choice of the loss function would lead to the median
- Proposals to replace mean and standard uncertainty as summary measures with…

Straying on forbidden grass…

Median and characteristic uncertainty, defined as

Characteristic uncertainty of X, denoted by $c(X)$, such that

$m(X) \pm 2c(X)$ is a 95 % coverage interval for X

Cox, M., O'Hagan, A. Meaningful expression of uncertainty in measurement. Accred Qual Assur 27, 19–37 (2022). https://doi.org/10.1007/s00769-021-01485-5

A similar proposal by A. Possolo and O. Bodnar

Straying on forbidden grass…

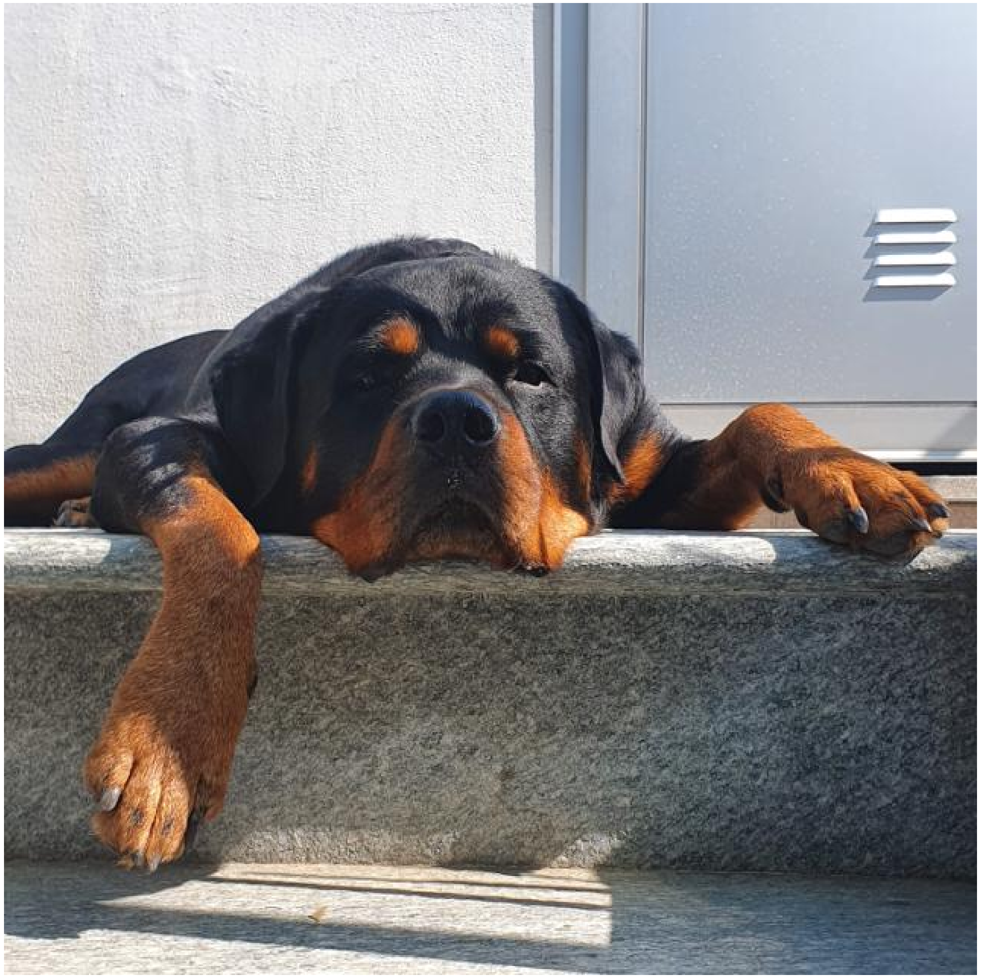
Advantages

- The median can be more representative than the mean
- Median and characteristic uncertainty can easily be calculated from the numerical approximation of the output (or posterior) PDF
- No longer need for expanded uncertainty, and easy calculation of a 95 % coverage interval

Main outstanding problems:

- A plausible counterpart for covariances
- Behaviour of these measures when propagated through a measurement model

My uncertain dog



Thank you