XIX IMEKO World Congress
Fundamental and Applied Metrology
September 6-11, 2009, Lisbon, Portugal

Accuracy, trueness, and precision: considerations based on the International Vocabulary of Metrology (VIM, 3rd Ed.) and related standards

R.Buccianti[1], M.Cibien[2], L.Mari[3], B.I.Rebaglia[4]

[1] CEI – Italian Electrotechnical Comittee, Milano, Italy
[2] UNI – Italian Organization for Standardization, Milano, Italy
[3] Università Cattaneo – LIUC, Castellanza (VA), Italy, lmari@liuc.it
[4] ITIA – CNR, Milano, Italy

Accuracy, trueness, and precision

Concepts of widespread usage in many metrological fields

but with non-trivial theoretical and operational differences in their meaning

A comparative analysis of these concepts as defined in

- ISO 5725: Accuracy (trueness and precision) of measurement methods and results
- ISO 3534: Statistics – Vocabulary and symbols
- VIM 3: International Vocabulary of Metrology – Basic and General Concepts and Associated Terms

Backgrounder

Let us assume that a sample of indications has been obtained by means of a measuring system in given measurement conditions

Several statistics can be computed on the sample, in particular:

- scale / dispersion statistics (p-quantiles, standard deviation)
- location / position statistics (median, mean)

They can be exploited to characterize the behavior of the measurement process, as resulting from the sample

A basic asymmetry:
the sample gives sufficient information only for scale statistics;
the value of a location statistic must be compared to a reference value

Reference values

- Theoretical values based on scientific principles
- Values obtained from the collaborative experimental work of a scientific or technical group, typically by a peer inter-laboratory comparison process
- Values obtained from the experimental work of some national or international recognized organization
- Values materialized by working standards, typically agreed on by customers / users and suppliers / manufacturers
- Values computed from series of previous observations of the same system

"Closeness of agreement"

All the 9 definitions we consider (3 terms x 3 documents) assume the concept of closeness of agreement as primitive, and apply it to both experimental values and reference values

Hence, by assuming:

c = closeness; *e* = experimental value; *r* = reference value

we will write, e.g.,

$$c(e, r)$$

to denote the closeness of agreement between *e* and *r*

Accuracy

- [5725-1] “closeness of agreement between a test result and the accepted reference value”
- [3534-2] “closeness of agreement between a test result or measurement result and the true value”
- [VIM 3] “closeness of agreement between a measured quantity value and a true quantity value of a measurand”

These definitions have the same form, $c(e, r)$, so that accuracy seems to be a **location** parameter, but radically differ about the reference value

- 5725: operational definition
- 3534: “in practice, the accepted reference value is substituted for the true value”
- VIM 3: “the concept ‘measurement accuracy’ is not a quantity and is not given a numerical quantity value”

More on accuracy

From the IEC 60050 series – International Electrotechnical Vocabulary:

- "closeness of the agreement between the result of a measurement and the conventionally true value of the measurand" {IEV, 394-40-35}
- "quality which characterizes the ability of a measuring instrument to provide an indicated value close to a true value of the measurand" {IEV, 311-06-08}
- "specified value of a parameter that represents the uncertainty in the measurement" {IEV, 415-05-12}

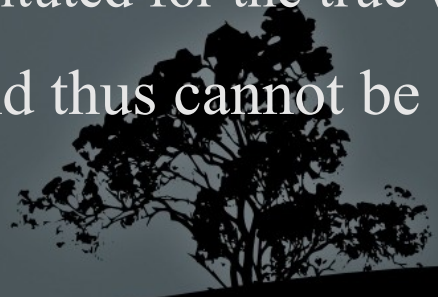
… Houston: we have a problem...

Trueness

- [5725-1] “closeness of agreement between the average value obtained from a large series of test results and an accepted reference value”
- [3534-2] “closeness of agreement between the expectation of a test result or a measurement result and a true value”
- [VIM 3] “closeness of agreement between the average of an infinite number of replicate measured quantity values and a reference quantity value”

These definitions have the same form, $c(e_{ave}, r)$, so that trueness seems to be a **location** parameter, but radically differ about both the experimental and the reference value

- 5725: operational definition
- 3534: “in practice, the accepted reference value is substituted for the true value”
- VIM 3: “measurement trueness is not a quantity and thus cannot be expressed numerically”



Precision

- [5725-1] “closeness of agreement between independent test results obtained under stipulated conditions”
- [3534-2] “closeness of agreement between independent test/measurement results obtained under stipulated conditions”
- [VIM 3] “closeness of agreement between indications or measured quantity values obtained by replicate measurements on the same or similar objects under specified conditions”

These definitions have the same form, $c(e_1, ..., e_n)$, so that trueness seems to be a **scale** parameter; they are substantially coincident and do not arise specific problems

A further problem

- [5725-1] “accuracy cannot be expressed in terms of bias or standard deviation only”
- [3534-2] “accuracy refers to a combination of trueness and precision”
- [VIM 3] accuracy “is related to both” trueness and precision

so that, contrary of what we have hypothesized in analyzing the definitions, accuracy is claimed to be not a location, but **an “overall” parameter**

For the discussion

Provided that this analysis is correct,

which strategy should the VIM adopt in a situation of inconsistency among different standards like the current one?

should it acknowledge the situation and include multiple, possibly mutually inconsistent, definitions?

or should it choose a single definition?

and in this case should the choice be made according to an assumed ”most common usage” or on the basis of an explicit conceptual model?

Some possible requirements for a conceptual model

These concepts should be defined so that:

- they specifically relate to measurement processes
- they do not imply any idealization of the measurement process, and therefore they are operative
- they maintain a clear distinction between the general (“qualitative”) concept and the possible parameters that quantify it (as the VIM 3 already does for precision: “measurement precision is usually expressed numerically by measures of imprecision, such as standard deviation, ...”)

and finally:

- accuracy maintains the role of overall concept