UNCERTAINTY IN MEASUREMENT: A SURVEY

Luca Mari
Libero Istituto Universitario C.Cattaneo
C.so Matteotti 22 , 21053 Castellanza (Va) - Italy
email: lmari@liuc.it

ISPRS Commission VI, Working Group 3

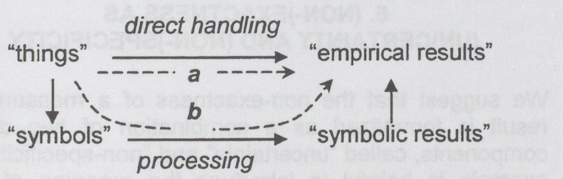
KEY WORD: Foundations of Measurement, Measurement Uncertainty

ABSTRACT:

The paper analyzes the concept of non-exactness of measurement results, by clearly distinguishing between (i) the way the results are expressed to make their uncertainty explicit; (ii) the way the chosen expression is interpreted as a suitable combination of non-specificity and uncertainty; (iii) the way the interpreted results are formally dealt with. In this perspective the merits and flaws of the ISO Guide to the expression of uncertainty in measurement are highlighted.

1. WHY NON-EXACTNESS IS AN ISSUE IN MEASUREMENT

Measurement is a means to set a bridge between the empirical world (to whom the measured thing belongs) and the linguistic world (to whom the measurement result belongs), aimed at enabling to faithfully re-interpret within the empirical world the information that has been obtained by handling symbols. The crucial point is the faithfulness of such a re-interpretation. In terms of the following diagram:



the issue is whether the procedures *a* and *b* would lead to the same result.

The fact is that the two worlds are inherently different. According to Bridgman, «there are certain human activities which apparently have perfect sharpness. The realm of mathematics and of logic is such a realm, par excellence. Here we have yes-no sharpness. But this yes-no sharpness is found only in the realm of things we say, as distinguished from the realm of things we do. Nothing that happens in the laboratory corresponds to the statement that a given point is either on a given line or it is not» (Bridgman, 1959).

The non-exactness (in the following the difference between non-exactness and uncertainty will be maintained and discussed) of measurement results accounts for such a distinction, although «by forcing the physical experience into the straight jacket of mathematics, with its yes-no sharpness, one is discarding an essential aspect of all physical experience and to that extent renouncing the possibility of exactly reproducing that experience. In this sense, the commitment of physics to the use of mathematics itself constitutes, paradoxically, a renunciation of the possibility of rigor» (Bridgman, 1959).

2. THE EXPRESSION OF NON-EXACT MEASUREMENT RESULTS

Taking into account the linguistic side of the problem, the first decision to be made is related to the form a measurement result should be given to make its non-exactness explicit. According to the ISO *Guide to the expression of uncertainty in measurement* (GUM) (ISO, 1993; a useful synthesis of the Guide can be found in Taylor, Kuyatt, 1997):

meas. result = <measurand value, uncertainty value>
= <*x*,*s*(*x*)>

where the first term of the couple could be obtained as the average of the population of the instrument readings and the second term as its estimated standard deviation.

Two notes in this regard:

* in presence of non-exactness the concepts of measurement result and measurand value cannot dealt with as equivalent: the measurement conveys information not only on the measurand value, but also on its uncertainty; in other terms, in this case a measurement result is not complete without the indication of a degree of uncertainty;

* different, and somehow more general, representations could be chosen, for example subsets, or fuzzy subsets, or probability distributions. All of these are more widely applicable than the representation suggested by the GUM, that can be employed only for algebraically strong scales (although the generality is understood, because of the applicability of the Chebyshev inequality, of expressing measurement results as couples <*x*,*s*(*x*)>,

being x and $s(x)$ the average and standard deviation of the (unknown) assumed probability distribution).

3. EVALUATION METHODS AND MEANINGS

The choice to express a non-exact measurement result as $<x,s(x)>$ still leaves open the decision about the evaluation methods that can be adopted to obtain such a result and the meaning to be attributed to it.

The GUM standpoint in reference to these decisions is peculiar. With respect to the evaluation methods, the GUM embodies a recommendation issued by the CIPM in 1981 (CIPM, 1981) and admits both statistical ("type A") and non-statistical ("type B") methods. The condition to make this pluralism operatively acceptable is that the suggested techniques to formally deal with the results are independent of the "type" of the evaluation method and therefore the same in both cases. From the conceptual point of view this position is an important step against the radical objectivism of some classical interpretations of measurement: some subjective information, in the form of "degrees of belief" (to quote the GUM) is present and required even in the case of an "objective" operation as measurement (Mari, Zingales, 1999).

Not so pluralistic is the position of the GUM in reference to the meaning of $s(x)$. Its basic interpretation is statistical, in terms of the standard deviation of the (possibly unknown) distribution of which x is the estimated average. The main application suggested by the GUM for this so-called "standard uncertainty" is to compute the "law of uncertainty propagation" (what is classically called "error propagation") through functional relations. Only for specific applications a further interpretation is recognized as useful, in which $s(x)$ (and more precisely $ks(x)$, being k a proportionality factor usually in the range [1,3]), in this case called "expanded uncertainty", is considered to express the half width of the interval of which x is the center point. This set-theoretic interpretation is however deemed as explicitly dependent of the statistical one and formally derived from it.

4. TWO CATEGORIES OF APPLICATIONS

The position of the GUM is conservative in this regard. A more general standpoint recognizes that the same result $<x,s(x)>$ could admit distinct interpretations in distinct applications. Given a set $\{<x_i,s(x_i)>\}$ of such measurement results, two basic categories of applications can be identified:

* "non-exact derived measurement": a quantity Y is known as analytically dependent of the quantities $X_1,\ldots,X_n$ through a function f, $Y=f(X_1,\ldots,X_n)$, and each $<x_i,s(x_i)>$ is the measurement result of a quantity X_i; the function f must be then somehow applied to the terms $<x_i,s(x_i)>$ to compute a measurement result $<y,s(y)>$ for Y;

* "non-exact measurement results comparison": all the $<x_i,s(x_i)>$ are repeated measurement results of the same quantity X, and must be compared to each other via a relation r, of which they are arguments, to establish whether such a relation holds among them or not.

Uncertainty propagation is clearly related to the first category, a case in which the statistical interpretation is plausibly the preferred one. The GUM suggests to compute y and $s(y)$ with distinct procedures, only in dependence on the terms x_i and $s(x_i)$ respectively:

$$y = f(x_1,\ldots x_n)$$

$$s(y) = [\textbf{Errore.}c_i^2\, s^2(x_i)]^{1/2}$$

being the "sensitivity coefficients" $c_i = \dfrac{\partial f}{\partial x_i}$ (in the case the quantities X_i are correlated further terms should be introduced).

The most important application in the second category is what could be called "non-exact equality", aimed at establishing whether two or more $<x_i,s(x_i)>$ can be considered undistinguishable with each other. The formal identity $<x_i,s(x_i)>=<x_j,s(x_j)>$, i.e. $x_i=x_j$ and $s(x_i)=s(x_j)$, is clearly a "too exact" criterion in this case, and different, more general, principles have been proposed, typically based on the set-theoretical interpretation of the terms $<x_i,s(x_i)>$. For example, if $<x_i,s(x_i)>$ is assumed as the interval $[x_i-s(x_i),x_i+s(x_i)]$ then two results could be judged "compatible" with each other in the case their intersection is non-null (cf. UNI, 1984).

It is worth to note that the GUM does not even mention this second category of applications. For important problems such as the definition of the procedures to compare national standards and express the results of the comparison an agreed position is still an open issue.

5. (NON-)EXACTNESS AS (UN)CERTAINTY AND (NON-)SPECIFICITY

We suggest that the non-exactness of a measurement result is formalized as a combination of two distinct components, called "uncertainty" and "non-specificity". An example is helpful to introduce the meaning of such concepts and their relations. Let us consider the following two statements:

A = "this is a 120-page book"

B = "this is a book"

aimed at expressing the knowledge of an observer on a given thing under examination. On the basis of the form of A and B two conclusions can be immediately drawn:

* A entails B: if A is true then also B must be true (in set-theoretical terms, A is a subset of B); therefore A is more specific than B;

* regardless of the particular uncertainty assignment chosen, A is at most as certain as B, and plausibly more uncertain than it

Hence the same formal expression, $<x,s(x)>$, admits two distinct, and actually opposite, meanings:

* the measurand value, the singleton x, is maximally specific but uncertain, with uncertainty $s(x)$;

* the measurand value, the interval $[x-s(x),x+s(x)]$, is not completely specific but considered certain.

The same empirical information, as obtained by the measurement, can be represented by suitably balancing the specificity and the certainty of the result (for example, if the statement A cannot considered certain on the basis of the available information, then

C = "this is a more-than-100-page book"

could be adopted, less specific but more certain than A. In the case the object is factually a 120-page book, then A would not be "truer" than C, but more specific, and therefore more informative, than it).

As a synthesis:

* measurement results are neither specific nor non-specific, and neither certain nor uncertain;

* they are neither singletons-with-uncertainty-degrees nor intervals;

* but they can interpreted in either way, in view of specific applications.

7. REFERENCES

Bridgman P.W., 1959. How much rigor is possible in physics? In: Henkin et al. (eds.), The Axiomatic method, North-Holland.

CIPM, BIPM, 1981. Proc. Com. Int. Poids et Mesures, 49, 8-9, 26 (in French) (also reported in the Introduction of the GUM).

ISO, et al., 1993. Guide to the expression of uncertainty in measurement. International Organization of Standardization, Geneva.

Mari, L., Zingales, G., 1999. Uncertainty in measurement science. In: Proceedings of the IMEKO TC7 Workshop, Kyoto.

Taylor N., Kuyatt C., 1997. Guidelines for Evaluating and Expressing the Uncertainty of NIST Measurement Results. NIST Technical Note 1297 (also available in the web: physics.nist.gov/Pubs/guidelines/contents.html).

UNI, 1984. Misure e misurazioni. Termini e definizioni fondamentali. Italian Standard UNI 4546.