A structured strategy of concept definition in measurement: the case of 'sensitivity'

2013 J. Phys.: Conf. Ser. 459 012040

(http://iopscience.iop.org/1742-6596/459/1/012040)

A structured strategy of concept definition in measurement: the case of 'sensitivity'

Luca Mari[1#], Arianna Mencattini[2]

[1] School of Engineering, Università Cattaneo – LIUC, Castellanza (VA), Italy
[2] University of Rome Tor Vergata, Italy

E-mail: lmari@liuc.it

Abstract. The paper emphasizes the importance that fundamental concepts in measurement science are defined according to a structured strategy, which provides both a general, qualitative characterization and a specific, type-related, quantitative definition. As a significant case, the concept 'sensitivity' is discussed and a definition for it proposed.

1. Introduction

The scope of measurement is broadening and its importance increasing: "an estimated 80 % [of the world trade] is affected by standards and regulations" and indeed according to various studies "the cost to producers and service providers of complying with "standards" can be 10 % of production costs" [1]. This generates new challenges not only towards the development of innovative (e.g., "smart") instrumentation, but also in the methodologies of data analysis, and at an even more fundamental level in the concept system by which measurement itself is interpreted and its results communicated. Measurement science has then become a moving target, as the *International vocabulary of metrology* (VIM3) [2] excellently witnesses: traditionally consolidated terms are changing their meaning (e.g., "measurement result") and new pivotal concepts are emerging (e.g., 'measurement model'). The VIM3 aim of being a well-structured and consistent concept system is critical and at the same time engaging. Indeed, the VIM3 is a guidance document "meant to be a common reference" for several different professional subjects and several different purposes [2: Scope]. With such an ambitious target, a significant issue is of properly balancing the strategy of concept definition so that definitions are at the same time encompassing and specific. This is particularly significant for those concepts which admit – or possibly require – a general (or "qualitative") characterization but also one or more ("quantitative") specifications. The way the three general features of measuring systems – measurement precision, measurement trueness, and measurement accuracy (whose definitions, by the way, in the VIM3 and in the standards ISO 3534 [3] and ISO 5725 [4] are distinctively different [5]) – are dealt with in the VIM3 is symptomatic on this matter: "measurement precision is usually expressed numerically by measures of imprecision..." [2: 2.15 N1]; "measurement trueness is not a quantity and thus cannot be expressed numerically, but measures for closeness of agreement are given..." [2: 2.14 N1]; "the concept 'measurement accuracy' is not a quantity and is not given a numerical quantity value." [2: 2.13 N1]. Hence, the situation is of

[#] To whom any correspondence should be addressed.
One of the authors is a member of the Joint Committee on Guides in Metrology (JCGM) Working Group 2 (VIM). The opinion expressed in this paper does not necessarily represent the view of this Working Group.

three general/"qualitative" concepts each of them with a claimed different treatment as for its specific/"quantitative" counterpart(s), without stated reasons justifying this position. The need of better constructing this general-specific derivation structure is acknowledged also by the *Guide to the expression of uncertainty in measurement* (GUM), when it states that "because of the lack of different words for this general concept of uncertainty and the specific quantities that provide quantitative measures of the concept, for example, the standard deviation, it is necessary to use the word "uncertainty" in these two different senses." [6: 2.2.1].

A possibly important case of multiple specific concepts under the same general concept has been triggered by the inclusion in the VIM3 of ordinal quantities [2: 1.26] and nominal properties [2: 1.30], complementing the traditional quantities based on measurement units. Even though not so well formalized yet as in the Stevens' theory of scale types [7], then extended in the so called representational theories of measurement [8], the systematic treatment of multiple property types [9] would be a significant step forward. We envisage a structure where a single general concept, say 'measurement uncertainty' can be coupled with a set of specific concepts, one for each type (and therefore 'uncertainty for quantities with unit', 'uncertainty for ordinal quantities', and so on). This would generate a very flexible framework, in which the concept system can be browsed according to several different viewpoints, such as "display the general concepts only" or "display everything which applies to ordinal quantities only".

It is in this perspective that the present work has been developed.

2. On sensitivity

An interesting case study of this possible framework is about the concept 'sensitivity', which is differently defined in different contexts and with different specifications (see, e.g., [10]: today – May 2013 – Electropedia lists 52 different lemmas containing the term "sensitivity"...). The VIM3 itself gives a definition – "quotient of the change in an indication of a measuring system and the corresponding change in a value of a quantity being measured" [2: 4.12] – whose general applicability, in particular in the case of ordinal quantities, is not explicitly discussed.

With the goal of making the many existing definitions comparable, and thus highlighting the possible set of meanings and their compatibility, let us assume:

Hypothesis 1 – Sensitivity is a property of an entity X modeled as a black box and under the only condition that the output $y(t_2)$ of X at the time t_2 causally depends on its input $u(t_1)$, $t_1 \leq t_2$.

X could be a measuring system but also a test, an algorithm, ... If X realizes an experimental process – the case of interest in measurement, where X is a measuring instrument – u and y are properties of the entity that interacts with X (or the environment itself) and X respectively. While the input-output behavior is then an empirical feature of X, related to the involved properties, its quantitative characterization is performed on the corresponding property values. For example, while the input of a spring is a force and its output is a length, the spring sensitivity is evaluated by comparing length values and force values. This distinction between properties and property values (and thus more specifically between quantities and quantity values: note that we are adopting the VIM3 position of considering quantities as specific properties) is crucial for appropriately modeling a measurement process but can be safely omitted here: we are discussing of the metrological characterization of X, not of a measurement by means of X. Hence for simplicity of notation in the following u and y will denote property values.

As the quoted VIM3 definition shows, Hypothesis 1 is immediately applicable to measuring instruments, by just assuming that y = value of the indication and u = value of the quantity being measured.

Hypothesis 1 admits that the analytical form of the input-output relation, i.e., the entity behavior, is not known, and does not impose any constraints on the types of the properties u and y.

As stated above, we are looking for both a general, "qualitative" concept and several type-related, "quantitative" concepts of sensitivity, so that there can be a nominal sensitivity for entities X whose input is a nominal property, an ordinal sensitivity for entities X whose input is an ordinal quantity, and so on, and under the condition that the type-related concepts are specifications of the general one.

Our second basic hypothesis is:

Hypothesis 2 – The sensitivity of X is a feature related to the ability of X to detect variations of its input.

Accordingly, our strategy is to assume that, independently of the type the input property u, X is characterized by its sensitivity by comparing the variation of the generated output y with the variation of the generating input u.

Hence, the concept of variation has to be revised and generalized so to make it applicable to all property evaluation types.

3. A type-aware characterization of 'variation'

In the case of quantities evaluated in interval or ratio scales the concept of variation for their values is intrinsically part of the scale itself, and therefore it can be exploited as is, being represented numerically by the difference of distinct values v_i in the scale:

$$\mathrm{var}(v_i, v_j) = |v_i - v_j|$$

or even the differential $\mathrm{d}v$ whenever continuously varying quantities are assumed.

Not so univocal is the concept for ordinal quantities, where the only information meaningfully available on the quantities is ranking and therefore their difference is undefined. On the other hand, variations in a totally ordered scale can be simply evaluated, e.g., by counting the number of values in the scale within the values corresponding to the quantities under consideration. If it is supposed that the scale index i runs monotonically with the scale order, i.e., the scale values are $v_1 < v_2 < v_3 < ...$, then trivially:

$$\mathrm{var}(v_i, v_j) = |i - j|$$

It can be easily shown that the so defined function $\mathrm{var}(v_i, v_j)$ is a metric.

In the case of nominal properties not even order can be exploited: two properties are either in the same class or in distinct classes, and therefore they are either associated to the same value or to distinct values, but neither distance nor order are meaningfully defined among such values. This suggests a concept of variation such as:

$$\mathrm{var}(v_i, v_j) = 1 - \delta_{i,j}$$

where the function:

$$\delta_{i,j} = 1 \text{ if } i = j \text{ and } = 0 \text{ otherwise}$$

is the Kronecker delta. It can be easily shown that also in this case $\mathrm{var}(v_i, v_j)$ is a metric.

Hence, for each property evaluation type at least one variation function $\mathrm{var}(v_i, v_j)$ is available that is applied to property values and represents information on the empirical variation of the corresponding properties. The fact that, independently of the type, $\mathrm{var}(v_i, v_j)$ is a metric, i.e., a function ranging in R, allows us to exploit it to define a concept of sensitivity which is at the same time qualitatively encompassing and quantitatively specifiable to each type.

4. A type-aware characterization of 'sensitivity'

Let us emphasize again the strategic importance of introducing both a general definition of 'sensitivity' and, on this basis, some type-related specific quantitative definitions. This provides two levels of access to information and reading of documentation:

– the general definition, at the qualitative level, offers to heterogeneous scientific groups, having to do more or less directly with measurement, a wide accessibility and easy interpretation of the concept;

– the specific definitions, at the quantitative level, are oriented to focused targets and can be customized according to specific theoretical assumptions or application needs.

General definition ("qualitative") – Sensitivity is a property of an entity modeled as a black box that characterizes its ability to produce a variation of its output in response to a variation of the received input.

NOTE – The entity could be, e.g., a measuring instrument, a process, a test, an algorithm.
NOTE – Since the entity has generally multiple input properties, a sensitivity can be defined for each input property.
NOTE – Sensitivity can depend on the input value: a constant sensitivity over the whole range of input values is generally a simplification hypothesis.
NOTE – Sensitivity can depend on time delays in the relation between input and output, and on transitory effects: a sensitivity independent of time is generally a simplification hypothesis.
NOTE – To assess the sensitivity of an entity two distinct inputs must be applied to it and the corresponding outputs must be recorded and compared.
NOTE – This definition is based on the generic concept 'variation', which admits different definitions in reference to different property evaluation types. Hence it does not provide any specific way to assign a value to sensitivity.

In all the following definitions, u and y are input and output property values respectively (for simplicity the time dependence is omitted).

General definition ("quantitative") – Sensitivity at the input value u is:

$$\mathrm{S}(u) = \frac{\mathrm{var}\left(y(u), y(u')\right)}{\mathrm{var}\left(u, u'\right)}$$

where var is a generic metric function and u' is an input value such that var(u, u') is the minimum positive value allowed for var(u, x), i.e., $u' = x$ where $\min_x \mathrm{var}(u, x) > 0$.

NOTE – A property u' that satisfies the given condition is not necessarily unique.

Specific definition 1 – Sensitivity at the input value u for an interval or ratio evaluation is:

$$\mathrm{S}(u) = \frac{dy(u)}{du} \quad \text{or} \quad = \frac{\Delta y(u)}{\Delta u} = \frac{y(u+\Delta u) - y(u)}{\Delta u}$$

for a continuous and a discrete quantity respectively.

Specific definition 2 – Sensitivity at the input value u_i for an ordinal evaluation is:

$$\mathrm{S}(u_i) = \frac{\mathrm{var}\left(y(u_i), y(u_{i+1})\right)}{\mathrm{var}\left(u_i, u_{i+1}\right)} = |k_0 - k_1|$$

where k_0 and k_1 are the ordinal indexes of $y(u_i)$ and $y(u_{i+1})$ respectively.

NOTE – In this case var(u_i, u_{i+1}) = 1.
NOTE – S(u_i) counts the number of ordinal positions between $y(u_i)$ and $y(u_{i+1})$.

Specific definition 3 – Sensitivity at the input value u_i and in reference to the input value u_j, $u_i \neq u_j$, for a nominal evaluation is:

$$\mathrm{S}_{u_j}(u_i) = \frac{\mathrm{var}\left(y(u_i), y(u_j)\right)}{\mathrm{var}\left(u_i, u_j\right)} = 1 - \delta_{k_0, k_1}$$

where δ is the Kronecker delta and k_0 and k_1 are the classificatory indexes of $y(u_i)$ and $y(u_j)$ respectively.

NOTE – In this case var(u_i, u_j) = 1.

NOTE – Because of the lack of structure on the sets of values, in this case the function S is actually biargumental.

NOTE – The possible values of $S_{u_j}(u_i)$ are either 0 or 1.

5. An example: computing sensitivity in a diagnostic test

The Breast Imaging-Reporting and Data System (BI-RADS) [11] is a quality assurance tool originally designed for use with mammography, published and trademarked by the American College of Radiology (ACR). The system is aimed at standardizing reporting, and is used by medical professionals to communicate a patient's risk of developing breast cancer. In particular, BI-RADS focuses on breast density, assumed as a factor positively correlated to risk of breast cancer, and defines a reference set to be used to evaluate Breast Tissue Composition (BTC). The set includes four elements, D1: the breast is almost entirely fat; D2: scattered fibroglandular densities (25-50%); D3: heterogeneously dense breast tissue (51-75%); D4: extremely dense (> 75% glandular), each element thus indicating the approximate percentage of dense tissue over the entire breast region. On this basis radiologists try to develop objective protocols to evaluate BTC in screening programs, where the protocol sensitivity has to be assessed for scientific purposes and specifically for comparing different protocols. BTC can be at first considered a nominal property by assuming that the four categories identified by the reference set D1–D4 only define a classification. Under this hypothesis the sensitivity of a protocol to evaluate BTC can be assessed, where the input is a mammographic image corresponding to a BTC in a given category $u_i = \mathrm{D}i$ and the output is a value $y_j = u(\mathrm{D}i) = \mathrm{D}j$ in the same set as obtained by the protocol (the evaluation procedure could be based on the independent examination of the same image by different expert radiologists, each of them being requested to identify one of the four categories, so that the output is the value corresponding to the most frequently chosen category).

Let us consider the nominal sensitivity of the protocol at the input value $u_1 = \mathrm{D1}$ (fatty tissue), hence by testing a BTC $u_i = \mathrm{D}i$, $i \neq 1$. Then, by applying results of Definition 3:

$$S_{\mathrm{D}i}(\mathrm{D}1) = \mathrm{var}(y(\mathrm{D}1), y(\mathrm{D}i)) = 1 - \delta_{1,i}$$

The sensitivity is equal to 1 if the protocol distinguishes the applied BTC from that of fatty tissue.

Let us now assume that BTC is an ordinal quantity, i.e., such that the relations D1 < D2 < D3 < D4 are empirically meaningful. The sensitivity at the input value D1, S(D1), is assessed by applying the protocol in the cases $u_1 = \mathrm{D1}$ and $u_2 = \mathrm{D2}$, so that $\mathrm{var}(u_1, u_2) = 1$. The possible outputs are such that S(D1) ranges from 0, if $y(\mathrm{D1}) = y(\mathrm{D2})$, i.e., the inputs are not distinguished with each other, to 3, in the case $y(\mathrm{D1}) = \mathrm{D1}$ and $y(\mathrm{D2}) = \mathrm{D4}$ (or vice versa $y(\mathrm{D1}) = \mathrm{D4}$ and $y(\mathrm{D2}) = \mathrm{D1}$).

6. Some conclusions

The sensitivity of an entity has been defined in the present paper in such a way that it is both qualitatively characterized and then specified quantitatively, in reference to each property evaluation type. The proposed definition is entirely based on the black box modeling of the entity: it assumes that two suitably chosen inputs are applied to the entity and the corresponding outputs are recorded and compared. An important consequence is that such a definition is calibration-independent: were it accepted, all definitions of 'sensitivity' that are instead formulated in terms of reference (or possibly true) values should be revised, because capturing some different concept. On the other hand, as characterized here 'sensitivity' is a so encompassing concept that it can be exploited also to define as specific cases, e.g., selectivity (as lack of sensitivity to properties other than the one intended to be evaluated) and repeatability (as lack of sensitivity to time).

The poor algebraic structure of nominal and ordinal evaluations leads to a characterization in terms of sensitivity that conveys a small amount of information, in particular a binary value in the nominal case. In this view, a possible generalization of the proposed concept builds upon the idea to perform a statistical analysis on the outputs obtained from a sample of inputs in repeatability conditions. Sensitivity would be then a function of a suitable statistic of the variations in such samples.

Journal of Physics: Conference Series **459** (2013) 012040 doi:10.1088/1742-6596/459/1/012040

We believe that the structural strategy used for 'sensitivity' may be generally adopted to produce an accessible, readable, and consistent concept system of measurement science fundamentals.

References

[1] BIPM 2007: Bureau International des Poids et Mesures, Evolving needs for metrology in trade, industry and society and the role of the BIPM [http://www.bipm.org/utils/en/pdf/Kaarls2007.pdf].

[2] JGCM 200:2012, International vocabulary of metrology – Basic and general concepts and associated terms (VIM), Joint Committee for Guides in Metrology, 2012 (2008 edition with minor corrections) [http://www.bipm.org/en/publications/guides/vim.html].

[3] ISO 3534, Statistics - Vocabulary and symbols, International Organization for Standardization, different publication dates for the various parts.

[4] ISO 5725, Accuracy (trueness and precision) of measurement methods and results, International Organization for Standardization, different publication dates for the various parts.

[5] R. Buccianti, M. Cibien, L. Mari, B. Rebaglia, Accuracy, trueness, and precision: considerations based on the International Vocabulary of Metrology (VIM, 3rd ed.) and related standards, proc. XIX IMEKO World Congress, Lisbon, Portugal, 2009.

[6] JGCM 100:2008, Evaluation of measurement data – Guide to the expression of uncertainty in measurement, Joint Committee for Guides in Metrology, 2008 (GUM, originally published in 1993) [http://www.bipm.org/en/publications/guides/gum.html].

[7] S. S. Stevens, On the theory of scales of measurement, Science, 103, 2684, 677–680, 1946.

[8] D. H. Krantz, R. D. Luce, P. Suppes, A. Tversky, Foundations of measurement, vols. 1–3, Academic Press, New York, 1971, 1989, 1990.

[9] A. Giordani, L. Mari, Property evaluation types, Measurement, 45, 437–452, 2012.

[10] IEC 60050, International Electrotechnical Vocabulary (IEV), International Electrotechnical Commission, different publication dates for the various parts [http://www.electropedia.org].

[11] American College of Radiology (ACR) Breast Imaging Reporting and Data System Atlas (BI-RADS® Atlas). Reston, Va: © American College of Radiology; 2003.