Probabilistic measurement modelling may overcome the opposition between the Bayesian and the frequentistic views

Giovanni Battista Rossi and Francesco Crenna
Università degli Studi di Genova –DIME/MEC, Laboratorio di Misure, Genova, Italy

E-mail: gb.rossi@dimec.unige.it.

Abstract. The probabilistic modelling of measurement systems is discussed, in regards to uncertainty evaluation and different models representing the same real system from different standpoints are compared. We suggest that proper probabilistic modelling directly yields evaluation of measurement uncertainty, without making any commitment to specific philosophical schools, such as the frequentistic and the Bayesian ones. We suggest that this model-based approach may have advantages in education and in the development of recommended practices in measurement.

1. Introduction
The Guide to the Expression of Uncertainty in Measurement was published in 1993 [3]1, in order to find an agreeable common way of evaluating and expressing uncertainty in measurement, and the requirement of accounting for both random variations and systematic effects that contribute to such uncertainty.

The GUM had an enormous impact in the field of measurement. On the theoretical side, it stimulated studies on the foundation and best methods for evaluating measurement uncertainty. Such studies were often concerned to the adoption of special visions of probability and of probabilistic inference such as the frequentistic and the Bayesian [4, 6]. Other studies were instead more concerned in addressing measurement modelling in general terms [9, 10].

On the application side, although the practice of actually evaluating uncertainty had a significant diffusion, difficulties were experienced, mainly due to a lack of practical guidelines for its application in the many scientific, technical and even social areas in which measurement is required.

After so many years from its original publication, a revision is now foreseen [12-16]. Yet current proposals seem to be closely related to the debate of the visions mentioned above. We suggest instead that the focus should rather be put to addressing measurement modelling, which would allow to overcome the traditional opposition between the Bayesians and the frequentists, and lead to a more user-oriented approach.

Therefore, in this communication, we consider the case of a measuring system which although being very simple, is general enough for our proposed, and we show how it can be properly modelled in different ways. Then we discuss such models, in respect also to the two mentioned competing visions. We suggest that the choice of a proper model is what really matters, since this naturally leads

1 References are listed in chronological order, to give a sense of the historical development of the subject.
to the solution by the application of the rules of probability calculus, without any additional philosophical commitment.

2. Different probabilistic models for the same measuring system
Consider the case of a linear measuring system for some quantity \( x \), for example a force (N), and producing an (instrument) indication \( y \), expressed for example as a voltage reading (mV), and having a sensitivity, \( k \) (mV/N).

Consider just two sources of uncertainty: a noise, \( v \), at the output of the voltmeter, and an uncertainty on the sensitivity, \( \delta k \), typically due to the calibration uncertainty for the device under consideration.

In order to obtain the final result, i.e., the measurement value, \( \hat{x} \), the (rough) instrument indication has to be scaled by dividing it by \( k \). Let us call observation the (physical) transformation operated by the measuring system, and restitution the (numerical) transformation that maps the indication into the measurement value. An illustration of this system is presented in figure 1, where \( c = k^1 \).

![Figure 1. The measuring system: model \( \mathcal{M}_1 \).](image)

In spite of its simplicity, this model allows us to discuss some features of measurement uncertainty. In fact \( v \) and \( \delta k \), although both constituting uncertainty source, are quite different entities. Indeed, \( v \) represents random noise and this variable is the typical component of the “classical” theory of errors and of the frequentistic approach, where probabilities are interpreted as expressing relative frequencies in the occurrence of events. In other words, probability as a description of “randomness” is perfectly suited here [5].

On the other hand, \( \delta k \) is assumed to express our “imperfect” knowledge of the value of the sensitivity, which is instead assumed to keep a constant value during the measurement process. Dealing with \( \delta k \) is typical of a Bayesian attitude, where probability is interpreted as expressing the state of knowledge of the experimenter, rather than the state of thing of the (real) system under consideration [2].

Considering the extreme cases, we could argue that if we had only uncertainty sources similar to \( v \), the classical, error-theory, frequentist approach would be perfectly suited. On the opposite side, if all the uncertainty sources would be similar to \( \delta k \), we may admit that a Bayesian, subjectivists, approach would be more appropriate. But the point is that in measurement both kinds of uncertainty sources typically coexist!

Let us then now discuss how this can be managed in three approaches:
- frequentistic,
- Bayesian, or
- model-based (the one we support).

Prior of entering into this discussion, let us first consider the model in figure 1, let us call it \( \mathcal{M}_1 \). Introducing the measurement error, \( e \),

\[
e = \hat{x} - x,
\]  

(1)
we obtain the sensitivity equation:

\[ e = \frac{x}{k} \delta k + cv, \]  

which yields, for standard uncertainty:

\[ u = \sigma_x = \sqrt{\frac{c^2 \sigma_x^2 + \frac{x^2 \sigma_{\delta k}^2}{k^2}}{k^2}} \approx \sqrt{\frac{c^2 \sigma_x^2 + \frac{x^2 \sigma_{\delta k}^2}{k^2}}{k^2}}, \]  

since the (unknown) value of the measurand, \( x \), can surely be approximated by the measurement value, \( \hat{x} \).

Yet there are alternative possibilities for modelling the same “situation” described in figure 1. One possibility is to leave \( v \) in the observation phase, since it is may be considered as an “objective” source of “randomness”, and to move \( \delta k \) to the restitution phase, since it is related to the state of knowledge of the measurer, who have an imperfect knowledge of the measuring system. Note that in figure 1, \( k \) is the “nominal” value of the sensitivity, i.e., the value we know for it, whilst the actual value is \( k + \delta k = k' \). Therefore, uncertainty arises, because we use \( k \) instead of \( k' \) in the restitution phase. Therefore, after considering that

\[ e = \frac{1}{k} \frac{1}{k - \delta k} = k'^{-1} \frac{1}{1 - \frac{\delta k}{k'}} \equiv k'^{-1} \left( 1 + \frac{\delta k}{k'} \right), \]

we obtain the model presented in figure 2. Let us call this model \( M_2 \).

![Figure 2. The measuring system: model \( M_2 \).](image)

Here the sensitivity equation reads:

\[ e = \frac{x}{k'} \delta k + \frac{1}{k'} v, \]  

and standard uncertainty can be evaluated by

\[ u = \sigma_x = \sqrt{k'^{-2} \sigma_x^2 + \frac{x^2 \sigma_{\delta k}^2}{k'^2}} \approx \sqrt{\frac{c^2 \sigma_x^2 + \frac{x^2 \sigma_{\delta k}^2}{k'^2}}{k'^2}}, \]  

since \( k' \) is, strictly speaking, unknown, and it can be replaced by \( k \). Therefore, model \( M_2 \) yields the same evaluation of standard uncertainty as model \( M_1 \).

Lastly, let us consider a third model that expresses uncertainty as expressing the incomplete knowledge of the measurer, in respect of the measurement process. In this perspective, uncertainty
sources are more appropriately referred to the restitution phase, as shown in figure 3, where model $M_i$ is presented.

![Figure 3. The measuring system: model $M_i$.](image)

Here a “correction” $\kappa = -v$ is virtually applied in order to compensate the effect of the noise $v$. But if such noise is assumed to have a null expected value, no compensation is actually apply, and the variable $\kappa$ is only considered in the uncertainty budget. That is to say: we are uncertain because we are unable to actually apply the compensation $\kappa$, as we would wish.

Here standard uncertainty is understood as the standard deviation of the measurement value, $\hat{x}$, for each given value of the indication $y$. We obtain:

$$u = \sigma_{\hat{x}|y} = \sqrt{k'^{-2}\sigma_k^2 + \frac{x^2\sigma_{\delta k}^2}{k'^2}} \approx \sqrt{c^2\sigma_k^2 + \frac{x^2\sigma_{\delta k}^2}{k'^2}}. \tag{6}$$

Again evaluated standard uncertainty is the same. Let us now discuss these three models.

3. Discussion of the models

Model $M_i$ is typical of the classic or frequentistic approach, since it ascribes all the uncertainty sources to the observation phase. Yet its application rises a problem: how can a fixed systematic effect such $\delta k$ be treated in probabilistic terms, if probability has to be related to some observable relative frequency in the occurrence of some event? In fact, the classical theory of (measurement) errors, assumed the absence of non-negligible systematic effects. Indeed, this is the main objection raised by the Bayesians to the frequentistic approach, in particular in the case of measurement.

Yet this objection can be circumvented, by referring to a kind of thought experiment.

In fact, we can consider the instrument we are using as randomly sampled by the set of instrument of the same kind, independently calibrated! If we repeated the observation not with the same instrument, but randomly sorting each time an instrument from the class, the variable $\delta k$ would also produce an observable variability, that corresponds to the standard deviation $\sigma_{\delta k}$ . Therefore, even in the case we use a single instrument, we can properly associate to it such dispersion, that it “inherit” from the class it belongs to.

On the opposite side, model $M_3$ is clearly related to a Bayesian approach.

This model also works, since it also yields the same result as the others, yet it also encounters some difficulties. For example, it treats the random effect by introducing a “virtual” correction that it is not actually applied: this is not such a straightforward argument, in our opinion. In particular, it is not less difficult to accept that the though experiment we have mentioned above, in relation to the frequentistic approach. Furthermore, standard uncertainty is now defined as a conditional quantity, but the conditioning term is never reported, as far as we now, in the final measurement result.
Lastly, model $\mathcal{M}_2$ is somewhat intermediate between the two, although probably closer to $\mathcal{M}_1$, for example in the way standard uncertainty is defined. It has the advantage of outlining the different way in which the two uncertainty sources affect the measurement process (one being a systematic effect, the other a source of random variations), but at the cost of a greater complexity.

In any case we can note that once a model has been adopted, it is possibly to evaluate the uncertainty by simply applying to the model the – universally accepted – rules of probability calculus. So where is the benefit of committing ourselves to one of the other philosophical approaches to probability? What this add to modelling? In the examples above we have seen that it just adds difficulties in their interpretation.

Our proposal is thus to approach measurement uncertainty evaluation by simply assuming a probabilistic model and then to drive the implied consequences.

If we agree on this, what model is to be preferred?

Let us that the problem is not dramatic, since at least in case we have considered all the models yield the same result. Yet there may be other reason for preferring one solution. In this regard, educational aspects also play a role.

4. Educational aspects

In University courses on measurement and instrumentation, modelling is essential: it is through models that students understand the functioning principles of devices that constitute measuring chains, and learn how to design them, for performing required tasks and to evaluate their performance [1, 7]. Therefore, when alternatives are possible, as it happens here, models should be chosen also accounting for their ability to effectively communicate the information.

In we compare, from this standpoint, the three models we have presented above, it is apparent how model $\mathcal{M}_3$ is by far the most difficult to understand.

Models $\mathcal{M}_1$ and $\mathcal{M}_2$ are both simpler and more straightforward than the previous one, and are quite similar to each other. The definition of the error, of the sensitivity function and of standard uncertainty as the standard deviation of the error, are clear and intuitive. Comparing them, we may note that $\mathcal{M}_2$ is perhaps more apt to highlight the difference between systematic effects and random variations, yet this is obtained at the expense of a higher complexity. For this reason, we use, in our courses, preferably model $\mathcal{M}_1$.

Concerning best practice guidelines, in our opinion what works better in education is, in our opinion likely to be more appropriate. For this reason we have recently made a proposal for revising the GUM in order to make it more user oriented, in which we have proposed model $\mathcal{M}_1$ as the most appropriate [16].

5. Conclusions

We have shown how the same measurement system can be correctly modelled in different ways, all of which allow us to obtain the same evaluation of standard uncertainty.

We suggest thus that uncertainty evaluation can be presented, in education and in recommended practice, with a focus on probabilistic modelling, without involving particular visions of probability such as the frequentist or the Bayesian, that may be confusing and do not help to solve the problem. The choice between different, although substantially equivalent models, should be made considering aspect of simplicity and effectiveness of communication.

Acknowledgements

This work was partially funded by the University of Genova through an Academic Research Program (FRA 2017).

References

[1]  Doeblin E 1983 Measurement Systems (Singapore: Mc Graw Hill)
[2] Press S J 1987 *Bayesian Statistics* (New York: Wiley)

[3] BIPM, IEC, IFCC, ISO, IUPAC, IUPAP, OIML 1993 *Guide to the Expression of Uncertainty in Measurement* Corrected and reprinted 1995 (Geneva, Switzerland: ISO) ISBN 92-67-10188-9

[4] Weise K and Wöger W 1993 A Bayesian theory of measurement uncertainty *Meas. Sci. Technol.* 4 pp 1-11

[5] Hacking I 2001 *An Introduction to Probability and Inductive Logic* (Cambridge, UK: Cambridge University Press)

[6] Lira I 2002 *Evaluating the Measurement Uncertainty* (Bristol: IOP)

[7] Bentley J P 2005 *Principles of Measurement Systems* (Harlow: Pearson-Prentice Hall)

[8] Morris A S and Langari R 2011 *Measurement and Instrumentation* (Academic Press)

[9] Frigerio A, Giordani A and Mari L 2010 Outline of a general model of measurement *Synthese* 7 pp 123-149

[10] Rossi G B 2012 Toward an interdisciplinary probabilistic theory of measurement *IEEE T. Instrum. Meas.* 61 pp 2097-106

[11] Rossi G B 2014 *Measurement and Probability* (Dordrecht: Springer)

[12] Bich W, Cox M and Michotte C 2016 Towards a new GUM – An update *Metrologia* 53 pp 149-59

[13] Lira I 2016 The GUM revision: The Bayesian view toward the expression of measurement uncertainty *Eur. J. Phys.* 37 025803

[14] White D R 2016 In pursuit of a fit-for-purpose uncertainty guide *Metrologia* 53 pp 107-124

[15] Possolo A and Pintar A L 2017 Plurality of Type A evaluations of uncertainty *Metrologia* 54 pp 617-632

[16] Rossi G B, Crenna F and Palazzo A 2018 A proposal for a new approach to the GUM *Proc. IEEE Int. Instrumentation and Measurement Technology Conf.* (I2MTC 2018), Houston, Texas, USA, May 14-17 - 2018, pp 1569-73