Contributions to the revision of the ‘Guide to the expression of uncertainty in measurement’

G A Kyriazis
Instituto Nacional de Metrologia, Qualidade e Tecnologia
Av. Nossa Senhora das Graças, 50 – Duque de Caxias – 25250-020 – RJ - Brazil
gakyriazis@inmetro.gov.br

Abstract. Some inconsistencies of the current version of the ‘Guide to the expression of uncertainty in measurement’ are discussed and suggestions to make this document consistent are commented. The paper is written taking into account the terminology of the third version of the ‘International vocabulary of metrology’.

1. Introduction
The Joint Committee for Guides in Metrology (JCGM) [1] was created in 1997 and is composed of representatives from the following organizations: Bureau International des Poids et Mesures (BIPM), International Electro-technical Commission (IEC), International Federation of Clinical Chemistry (IFCC), International Laboratory Accreditation Cooperation (ILAC), International Organization for Standardization (ISO), International Union of Pure and Applied Chemistry (IUPAC), International Union of Pure and Applied Physics (IUPAP) and International Organization of Legal Metrology (OIML). This committee is responsible for updating the ‘International vocabulary of metrology’ (VIM) and the ‘Guide to the expression of uncertainty in measurement’ (GUM) [2] (and its supplements). The VIM had its third version (VIM3), which was published in 2008 [3] (Portuguese-Brazilian translation in 2012 [4]), strongly influenced by the GUM uncertainty approach. Inconsistencies in the GUM text have been pointed out since its elaboration and publication in 1993. The concepts of Bayesian probability theory provide a consistent method for the quantitative reasoning under incomplete information. Suggestions to make the GUM consistent with such theory were proposed in 2003 [5], but at that time it was thought that it would be better to defer a revision of that document to a later time as a great effort had been done to disseminate its concepts. Nevertheless, in 2006, the JCGM decided on the elaboration of a consistent set of GUM supplements [6]. In 2012, the JCGM finally decided to revise the GUM to make it consistent both internally and with relation to its supplements [7].

The main inconsistencies of the current GUM are commented in section 2. The suggestions proposed in [5], which are endorsed by this author, are presented again here in section 3 (with small additions) for convenience of the reader.

The terminology of VIM3 is fully employed here. For convenience of the reader, the terms of that document are marked in italics the first time they appear in the text.

2. Inconsistencies in GUM
The initial step in an evaluation of measurement uncertainty (VIM3, 2.26) is to establish the measurement model (VIM3, 2.48). The latter, as defined in VIM3, is the ‘mathematical relation among
all quantities (VIM3, 1.1) known to be involved in a measurement (VIM3, 2.1)’. The generic form of the model is \( h(Y, X_1, ..., X_n) = 0 \), where \( Y \), the output quantity (VIM3, 2.51), is the measurand (VIM3, 2.3), whose value should be inferred based on the information about the input quantities (VIM3, 2.50) \( X_1, ..., X_n \). The latter can be: (a) the expected value(s) of the frequency distribution(s) of the (several sets of) indications (VIM3, 4.1), (b) the corrections (VIM3, 2.53) and (c) the influence quantities (VIM3, 2.52). In general, the measurement model can be expressed so that the output quantity becomes explicit, that is \( Y = f(X_1, ..., X_n) \), where the function \( f \) is the measurement function (VIM3, 2.49). This measurement function is employed by the GUM and its supplements in the evaluation of measurement uncertainty.

Note that indications are neither input quantities as informed in VIM3, 2.50, Note 2 nor measured values (VIM3, 2.10) as alluded in VIM3, 2.28. Indications are ‘quantity values (VIM3, 1.19) provided by a measuring instrument (VIM, 3.1) or a measuring system (VIM, 3.2)’. Once they are registered, they are to be considered as fixed and known. They are used to infer the expected value (assumed fixed and unknown) of the frequency distribution (assumed known) of the indications. The GUM Supplement 1 (GUMS1) [8] presents the inference results for two frequency distributions: one continuous (Gaussian) (GUMS1, 6.4.9) and one discrete (Poisson) (GUMS1, 6.4.11). It also presents several probability distributions for the values of the other input quantities (about which information cannot be obtained, or one does not desire it to be obtained, from indications), that can be used according to the information available about them.

The current GUM prescribes two methods of uncertainty evaluation: one that takes into account the availability of indications, Type A evaluation (VIM3, 2.28), and the other, Type B evaluation (VIM3, 2.29), which is applicable in the absence of those. The mathematical expressions used in Type A evaluation, such as those described in the current GUM, are derived from conventional statistics, whilst the GUM interprets the combined standard uncertainty (VIM3, 2.31) from a Bayesian viewpoint. This is an inconsistency that should be corrected in the future revision of the GUM. Observe that this inconsistency does not occur in GUMS1 where the classification in Type A and Type B evaluations is not necessary (GUMS1, 5.11.4). On the other hand, one may continue with the distinction between the two evaluation methods, if one wishes to do so, as they are in fact different.

VIM3 defines measurement result (VIM3, 2.9) as the ‘set of quantity values being attributed to a measurand together with any other available relevant information’. The measurement result may be then expressed as (it is assumed here that the definitional uncertainty (VIM3, 2.27) is negligible in comparison with the other uncertainties involved):

- a probability distribution for the (essentially unique) true value (VIM3, 2.11) of the measurand based on all information available from the measurement. A numerical approximation of this distribution can be obtained by the method of propagation of distributions described in GUMS1;

- an interval of possible values (coverage interval (VIM3, 2.36)) within which it is believed that the (essentially unique) true value of the measurand is contained with a given coverage probability (VIM3, 2.37). GUMS1 describes how to calculate coverage intervals for given probabilities;

- a measured value (VIM3, 2.10) and its associated standard uncertainty (VIM3, 2.30). The measured value of the measurand is interpreted as the expected value of the probability distribution for the value of the measurand and the standard uncertainty associated with the measured value is interpreted as the standard deviation of this distribution. GUMS1 describes how to calculate the expected value and standard deviation of the numerical approximation of the probability distribution. If the model is linear or linearized, it is also possible to use the method described in GUM to obtain the expected value and variance of that distribution (note that the GUM does not allow one to specify completely the distribution) from the expected values and variances of the probability distributions for the input quantities.
Consider the measured value \( y \) and the combined standard uncertainty \( u(y) \). It can be demonstrated that the minimum and maximum coverage probabilities of the interval \([y \pm 2u(y)]\) are 75\% and 95\%, respectively, for any distribution with expected value \( y \) and standard deviation \( u(y) \) [5]. Therefore, the interval \([y \pm 2u(y)]\) covers a great fraction of the probability distribution represented by \( y \) and \( u(y) \). As the GUM procedure does not yield a complete specification of the distribution represented by \( y \) and \( u(y) \), the concept of minimum coverage probability of the interval \([y \pm 2u(y)]\) is in agreement with the GUM. In this case, \( 2u(y) \) is the expanded measurement uncertainty (VIM3, 2.35).

The current GUM prescribes two methods for the evaluation of the expanded measurement uncertainty. Though relegated to an appendix (GUM, Appendix G), such methods have systematically been adopted by the metrology (VIM3, 2.2) community when issuing calibration (VIM3, 2.39) certificates. The first one presupposes that the probability distribution for the value of the measurand can be approximated by a Gaussian. The second presupposes that the distribution can be approximated by a Student-\( t \) and prescribes the use of the Welch-Satterthwaite formula for the calculation of the corresponding degrees of freedom. The inconsistency of the last method is related to the fact that this method is not compatible with quantities whose measurement uncertainties (associated with their measured values) were obtained from a Type B evaluation. This is an inconsistency that should be corrected in the future revision of the GUM. Observe that this inconsistency does not occur in GUMS1 where the numerical approximation of the probability distribution is known and the expanded measurement uncertainty can be directly calculated from it.

As informed in GUMS1, section 6.4.9.4, Note 2: ‘In the Bayesian context of this Supplement, concepts such as the reliability, or the uncertainty, of an uncertainty are not necessary. Accordingly, the degrees of freedom in a Type A evaluation of uncertainty is no longer viewed as a measure of reliability, and the degrees of freedom in a Type B evaluation do not exist.’ The future version of the GUM should be correspondingly revised.

One drawback of the evaluation method based on the measurement function, as described in GUM and its supplements, is that prior information about the measurand cannot be used in a new evaluation of that same measurand (recalibration). On the other hand, the method allows for prior information about the expected value(s) of the frequency distribution(s) of the (several sets of) indications to be included in the new evaluation. It also permits the incorporation of the definitional uncertainty in the evaluation.

3. Steps of the future GUM

The steps of the evaluation of uncertainty in measurement that the author recommends be adopted in the revision of the GUM to make it consistent are set below. Such steps are in [5] and are presented here with small additions using the VIM3 terminology for convenience of the reader.

Step 1. Describe the measurement method (VIM3, 2.5) as a measurement function \( Y = f(X_1, \ldots, X_n) \), where the input quantities \( X_1, \ldots, X_n \) may have their own measurement functions and be evaluated by following the same steps described here. All input and output quantities are variables with probability distributions representing states of knowledge. Such distributions are obtained from the information available in the measurement as described in GUMS1.

Step 2. Determine the measured values \( x_1, \ldots, x_n \) of the input quantities \( X_1, \ldots, X_n \), respectively. All the measured values are interpreted as expected values of the probability distributions for the input quantities.

Step 3. Determine the standard uncertainties \( u(x_1), \ldots, u(x_n) \) associated with the measured values of the input quantities \( X_1, \ldots, X_n \), respectively. All the standard uncertainties are interpreted as standard deviations of the probability distributions for the input quantities.

Step 4. Evaluate if the probability distributions for each pair \( X_i \) and \( X_j \) of the input quantities are logically dependent and determine the correlation coefficients \( r(x_i, x_j) \) for those pairs of the input variables that may be significantly correlated, for \( i, j = 1, \ldots, n, i < j \).
Step 5. Determine the measured value \( y \) of the output quantity substituting all measured values \( x_1, \ldots, x_n \) of the input quantities in the measurement function \( Y = f(X_1, \ldots, X_n) \). Thus, \( y = f(x_1, \ldots, x_n) \).

Step 6. Determine the combined standard uncertainty \( u(y) \) associated with the measured value \( y \) by propagating the standard uncertainties \( u(x_1), \ldots, u(x_n) \) and the correlation coefficients \( r(x_1, x_2), \ldots, r(x_{n-1}, x_n) \). A first-order Taylor series approximation of the measurement function \( Y = f(X_1, \ldots, X_n) \) provides the following law of propagation of uncertainties:

\[
u^2(y) = \sum_i c_i^2 u^2(x_i) + 2 \sum_{i<j} c_i c_j u(x_i) u(x_j) r(x_i, x_j), \tag{1}\]

where \( c_1, \ldots, c_n \) are the sensitivity coefficients or partial derivatives of \( Y \) with respect to \( X_1, \ldots, X_n \) evaluated at \( x_1, \ldots, x_n \), respectively. The product \( u(x_i) u(x_j) r(x_i, x_j) \) is equal to the covariance \( u(x_i, x_j) \) for \( i, j = 1, \ldots, n, i < j \).

Step 7. If it is necessary to express the measurement result as a coverage interval, multiply the combined standard uncertainty \( u(y) \) by a coverage factor (VIM3, 2.38) \( k \) to obtain the coverage interval \([y \pm ku(y)]\). Generally, \( k = 2 \). Use values of \( k \) other than 2 in special cases only.

Step 8. Report the measurement result as the measured value \( y \) together with its associated combined standard uncertainty \( u(y) \) or as the coverage interval \([y \pm ku(y)]\). Describe how \( y \) and \( u(y) \) were obtained. Describe how \( k \) was chosen when \( k \) is not 2.

Note that if the evaluation is performed using only the method of propagation of uncertainties described in GUM, it will be no more necessary to report the coverage probability of the interval when the coverage factor is 2. In addition, it will not be needed anymore to use the Welch-Satterthwaite formula to calculate effective degrees of freedom.

References

[1] Joint Committee for Guides in Metrology, http://www.bipm.org/en/committees/jc/jcgm/ (accessed 13.11.13).

[2] BIPM, IEC, IFCC, ILAC, ISO, IUPAC, IUPAP and OIML, Evaluation of measurement data – Guide to the expression of uncertainty in measurement, GUM 1995 with minor corrections, Joint Committee for Guides in Metrology, JCGM 100, 2008, http://www.bipm.org/utils/common/documents/jcgm/JCGM_100_2008_E.pdf (accessed 13.11.13).

[3] BIPM, IEC, IFCC, ILAC, ISO, IUPAC, IUPAP and OIML, International vocabulary of metrology – Basic and general concepts and associated terms, JCGM 200, 2008, http://www.bipm.org/utils/common/documents/jcgm/JCGM_200_2008.pdf (accessed 13.11.13).

[4] BIPM, IEC, IFCC, ILAC, ISO, IUPAC, IUPAP and OIML, Vocabulário internacional de metrologia – Conceitos fundamentais e gerais e termos associados, 1ª edição luso-brasileira, 2012, http://www.inmetro.gov.br/inovacao/publicacoes.asp (accessed 13.11.13).

[5] Kacker R and Jones A 2003 Metrologia 40 235-48.

[6] Bich W, Cox M G and Harris P M 2006 Metrologia 43 S161-S166.

[7] Bich W, Cox M G, Dybkaer R, Elster C, Estler W T, Hibbert B, Imai H, Kool W, Michotte C, Nielsen L, Pendrill L, Sidney S, van der Veen A M H and Wöger W 2012 Metrologia 49 702

[8] BIPM, IEC, IFCC, ILAC, ISO, IUPAC, IUPAP and OIML, Evaluation of measurement data – Supplement 1 to the ‘Guide to the expression of uncertainty in measurement’, Propagation of distributions using a Monte Carlo method, Joint Committee for Guides in Metrology, JCGM 101, 2008, http://www.bipm.org/utils/common/documents/jcgm/JCGM_101_2008_E.pdf (accessed 13.11.13).