GUIDE 98-1

Uncertainty of measurement —
Part 1:
Introduction to the expression of uncertainty in measurement

First edition 2009

ISO copyright office
Case postale 56 • CH-1211 Geneva 20
Tel. + 41 22 749 01 11
Fax + 41 22 749 09 47
E-mail copyright@iso.org
Web www.iso.org

Published in Switzerland

ISO/IEC Foreword

ISO (the International Organization for Standardization) and IEC (the International Electrotechnical Commission) form the specialized system for worldwide standardization. National bodies that are members of ISO or IEC participate in the development of International Standards through technical committees established by the respective organization to deal with particular fields of technical activity. ISO and IEC technical committees collaborate in fields of mutual interest. Other international organizations, governmental and non-governmental, in liaison with ISO and IEC, also take part in the work.

Draft Guides adopted by the responsible Committee or Group are circulated to the member bodies for voting. Publication as a Guide requires approval by at least 75 % of the member bodies casting a vote.

Attention is drawn to the possibility that some of the elements of this document may be the subject of patent rights. ISO shall not be held responsible for identifying any or all such patent rights.

ISO/IEC Guide 98-1 was prepared by Working Group 1 of the Joint Committee for Guides in Metrology (as JCGM 104:2009), and was adopted by the national bodies of ISO and IEC.

ISO/IEC Guide 98 consists of the following parts, under the general title *Uncertainty of measurement*:

— *Part 1: Introduction to the expression of uncertainty in measurement*

— *Part 3: Guide to the expression of uncertainty in measurement (GUM:1995)*

The following parts are planned:

— *Part 2: Concepts and basic principles*

— *Part 4: Role of measurement uncertainty in conformity assessment*

— *Part 5: Applications of the least-squares method*

ISO/IEC Guide 98-3 has one supplement.

— *Supplement 1: Propagation of distributions using a Monte Carlo method*

The following supplements to ISO/IEC Guide 98-3 are planned:

— *Supplement 2: Models with any number of output quantities*

— *Supplement 3: Modelling*

Given that ISO/IEC Guide 98-1:2009 is identical in content to JCGM 104:2009, the decimal symbol is a point on the line in the English version.

Annex ZZ has been appended to provide a list of corresponding ISO/IEC Guides and JCGM guidance documents for which equivalents are not given in the text.

Joint Committee for Guides in Metrology

JCGM
104
2009

Evaluation of measurement data — An introduction to the "Guide to the expression of uncertainty in measurement" and related documents

Évaluation des données de mesure – Une introduction au "Guide pour l'expression de l'incertitude de mesure" et aux documents qui le concernent

Copyright

Even if electronic versions are available free of charge on the website of one or more of the JCGM member organizations, economic and moral copyrights related to all JCGM publications are internationally protected. The JCGM does not, without its written authorisation, permit third parties to rewrite or re-brand issues, to sell copies to the public, or to broadcast or use on-line its publications. Equally, the JCGM also objects to distortion, augmentation or mutilation of its publications, including its titles, slogans and logos, and those of its member organizations.

Official versions and translations

The only official versions of documents are those published by the JCGM, in their original languages.

The JCGM's publications may be translated into languages other than those in which the documents were originally published by the JCGM. Permission must be obtained from the JCGM before a translation can be made. All translations should respect the original and official format of the formulae and units (without any conversion to other formulae or units), and contain the following statement (to be translated into the chosen language):

> All JCGM's products are internationally protected by copyright. This translation of the original JCGM document has been produced with the permission of the JCGM. The JCGM retains full internationally protected copyright on the design and content of this document and on the JCGM's titles, slogan and logos. The member organizations of the JCGM also retain full internationally protected right on their titles, slogans and logos included in the JCGM's publications. The only official version is the document published by the JCGM, in the original languages.

The JCGM does not accept any liability for the relevance, accuracy, completeness or quality of the information and materials offered in any translation. A copy of the translation shall be provided to the JCGM at the time of publication.

Reproduction

The JCGM's publications may be reproduced, provided written permission has been granted by the JCGM. A sample of any reproduced document shall be provided to the JCGM at the time of reproduction and contain the following statement:

> This document is reproduced with the permission of the JCGM, which retains full internationally protected copyright on the design and content of this document and on the JCGM's titles, slogans and logos. The member organizations of the JCGM also retain full internationally protected right on their titles, slogans and logos included in the JCGM's publications. The only official versions are the original versions of the documents published by the JCGM.

Disclaimer

The JCGM and its member organizations have published this document to enhance access to information about metrology. They endeavor to update it on a regular basis, but cannot guarantee the accuracy at all times and shall not be responsible for any direct or indirect damage that may result from its use. Any reference to commercial products of any kind (including but not restricted to any software, data or hardware) or links to websites, over which the JCGM and its member organizations have no control and for which they assume no responsibility, does not imply any approval, endorsement or recommendation by the JCGM and its member organizations.

Contents

Foreword

In 1997 a Joint Committee for Guides in Metrology (JCGM), chaired by the Director of the BIPM, was created by the seven international organizations that had originally in 1993 prepared the 'Guide to the expression of uncertainty in measurement' (GUM) and the 'International vocabulary of metrology – basic and general concepts and associated terms' (VIM). The JCGM assumed responsibility for these two documents from the ISO Technical Advisory Group 4 (TAG4).

The Joint Committee is formed by the BIPM with the International Electrotechnical Commission (IEC), the International Federation of Clinical Chemistry and Laboratory Medicine (IFCC), the International Laboratory Accreditation Cooperation (ILAC), the International Organization for Standardization (ISO), the International Union of Pure and Applied Chemistry (IUPAC), the International Union of Pure and Applied Physics (IUPAP), and the International Organization of Legal Metrology (OIML).

JCGM has two Working Groups. Working Group 1, 'Expression of uncertainty in measurement', has the task to promote the use of the GUM and to prepare Supplements and other documents for its broad application. Working Group 2, 'Working Group on International vocabulary of basic and general terms in metrology (VIM)', has the task to revise and promote the use of the VIM. For further information on the activity of the JCGM, see www.bipm.org.

The present document has been prepared by Working Group 1 of the JCGM, and has benefited from detailed reviews undertaken by member organizations of the JCGM.

This document constitutes one part in a series of JCGM documents under the generic heading *Evaluation of measurement data*. The parts in the series are

— JCGM 100:2008. Evaluation of measurement data – Guide to the expression of uncertainty in measurement (GUM) (see clause 2),

— JCGM 101:2008. Evaluation of measurement data – Supplement 1 to the "Guide to the expression of uncertainty in measurement" – Propagation of distributions using a Monte Carlo method (see clause 2),

— JCGM 102. Evaluation of measurement data – Supplement 2 to the "Guide to the expression of uncertainty in measurement" – Models with any number of output quantities,

— JCGM 103. Evaluation of measurement data – Supplement 3 to the "Guide to the expression of uncertainty in measurement" – Modelling,

— JCGM 104. Evaluation of measurement data – An introduction to the "Guide to the expression of uncertainty in measurement" and related documents [this document],

— JCGM 105. Evaluation of measurement data – Concepts and basic principles,

— JCGM 106. Evaluation of measurement data – The role of measurement uncertainty in conformity assessment, and

— JCGM 107. Evaluation of measurement data – Applications of the least-squares method.

Introduction

A statement of measurement uncertainty is indispensable in judging the fitness for purpose of a measured quantity value. At the greengrocery store the customer would be content if, when buying a kilogram of fruit, the scales gave a value within, say, 2 grams of the fruit's actual weight. However, the dimensions of components of the gyroscopes within the inertial navigation systems of commercial aircraft are checked by measurement to parts in a million for correct functioning.

Measurement uncertainty is a general concept associated with any measurement and can be used in professional decision processes as well as judging attributes in many domains, both theoretical and experimental. As the tolerances applied in industrial production become more demanding, the role of measurement uncertainty becomes more important when assessing conformity to these tolerances. Measurement uncertainty plays a central role in quality assessment and quality standards.

Measurement is present in almost every human activity, including but not limited to industrial, commercial, scientific, healthcare, safety and environmental. Measurement helps the decision process in all these activities. Measurement uncertainty enables users of a measured quantity value to make comparisons, in the context of conformity assessment, to obtain the probability of making an incorrect decision based on the measurement, and to manage the consequential risks.

This document serves as an introduction to measurement uncertainty, the GUM and the related documents indicated in the Foreword. A probabilistic basis for uncertainty evaluation is used. Annex A gives acronyms and initialisms used in this document.

In future editions of JCGM 200 (VIM) it is intended to make a clear distinction between the use of the term *error* as a quantity and as a quantity value. The same statement applies to the term *indication*. In the current document such a distinction is made. JCGM 200:2008 does not distinguish explicitly between these uses.

Evaluation of measurement data — An introduction to the 'Guide to the expression of uncertainty in measurement' and related documents

1 Scope

The Joint Committee for Guides in Metrology (JCGM) has prepared this document to promote the sound evaluation of measurement uncertainty through the use of the GUM (see clause 2) and to provide an introduction to the GUM Supplements and other documents JCGM is producing: JCGM 101:2008 (see clause 2) and references [3, 4, 5, 6, 7].

As in the GUM, this document is primarily concerned with the expression of uncertainty relating to the measurement of a well-defined quantity—*the measurand* [JCGM 200:2008 (VIM) 2.3]—that can be characterized by *an essentially unique true value* [JCGM 200:2008 (VIM) 2.11 NOTE 3]. The GUM provides a rationale for not using the term 'true', but this term will be kept in this document when there is otherwise a possibility for ambiguity or confusion.

The purpose of the GUM Supplements and the other documents produced by the JCGM is to help with the interpretation of the GUM and enhance its application. The GUM Supplements and the other documents are together intended to have a scope that is considerably broader than that of the GUM.

This document introduces measurement uncertainty, the GUM, and the GUM Supplements and other documents that support the GUM. It is directed predominantly at the measurement of quantities that can be characterized by continuous variables such as length, temperature, time, and amount of substance.

This introductory document is aimed at the following, including but not limited to

— scientific activities and disciplines in general,

— industrial activities and disciplines in general,

— calibration, testing and inspection laboratories in industry, and laboratories such as those concerned with health, safety and environment, and

— evaluation and accreditation bodies.

It is hoped that it will also be useful to designers, because a product specification that takes better account of inspection requirements (and the associated measurement) can result in less stringent manufacturing requirements. It is also directed at academia, with the hope that more university departments will include modules on measurement uncertainty evaluation in their courses. As a result, a new generation of students would be better armed to understand and provide uncertainty statements associated with measured quantity values, and thus gain an improved appreciation of measurement.

This introductory document, the GUM, the GUM Supplements and the other documents should be used in conjunction with the 'International Vocabulary of Metrology—Basic and general concepts and associated terms' and all three parts of ISO 3534 cited in clause 2, which define statistical terms (used in statistics and probability, including applied statistics and design of experiments), and express them in a conceptual framework in accordance with normative terminology practice. The last consideration relates to the fact that the theoretical background of evaluation of measurement data and evaluation of the uncertainty of measurement is supported by mathematical statistics and probability.

2 Normative references

The following referenced documents are indispensable for the application of this document. For dated references, only the edition cited applies. For undated references, the latest edition of the referenced document (including any amendments) applies.

JCGM 100:2008. Evaluation of measurement data — Guide to the expression of uncertainty in measurement (GUM). Joint Committee for Guides in Metrology.

JCGM 101:2008. 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 200:2008. International Vocabulary of Metrology—Basic and general concepts and associated terms, 3rd Edition. Joint Committee for Guides in Metrology

ISO 3534-1:2006. Statistics – Vocabulary and symbols – Part 1: General statistical terms and terms used in probability.

ISO 3534-2:2006. Statistics – Vocabulary and symbols – Part 2: Applied statistics.

ISO 3534-3:1999. Statistics – Vocabulary and symbols – Part 3: Design of experiments.

3 What is measurement uncertainty?

3.1 The purpose of measurement is to provide information about a quantity of interest—a *measurand* [JCGM 200:2008 (VIM) 2.3]. The measurand might be the volume of a vessel, the potential difference between the terminals of a battery, or the mass concentration of lead in a flask of water.

3.2 No measurement is exact. When a quantity is measured, the outcome depends on the measuring system [JCGM 200:2008 (VIM) 3.2], the measurement procedure, the skill of the operator, the environment, and other effects [1]. Even if the quantity were to be measured several times, in the same way and in the same circumstances, a different *indication value* [JCGM 200:2008 (VIM) 4.1] (measured quantity value [JCGM 200:2008 (VIM) 2.10]) would in general be obtained each time, assuming that the measuring system has sufficient resolution to distinguish between the indication values. Such indication values are regarded as instances of an indication quantity.

3.3 The *dispersion* of the indication values would relate to how well the measurement is made. Their *average* would provide an *estimate* [ISO 3534-1:2006 1.31] of the *true quantity value* [JCGM 200:2008 (VIM) 2.11] that generally would be more reliable than an individual indication value. The dispersion and the number of indication values would provide information relating to the average value as an estimate of the true quantity value. However, this information would not generally be adequate.

3.4 The measuring system may provide indication values that are not dispersed about the true quantity value, but about some value offset from it. The difference between the offset value and the true quantity value is sometimes called the *systematic error value* [JCGM 200:2008 (VIM) 2.17]. Take the domestic bathroom scales. Suppose they are not set to show zero when there is nobody on the scales, but to show some value offset from zero. Then, no matter how many times the person’s mass were re-measured, the effect of this offset would be inherently present in the average of the indication values. In general, a systematic error, regarded as a quantity, is a component of error that remains constant or depends in a specific manner on some other quantity.

3.5 There are two types of measurement error quantity, *systematic* and *random* [JCGM 200:2008 (VIM) 2.19]. A systematic error (an estimate of which is known as a *measurement bias* [JCGM 200:2008 (VIM) 2.18]) is associated with the fact that a measured quantity value contains an offset. A random error is associated with the fact that when a measurement is repeated it will generally provide a measured quantity value that is different from the previous value. It is random in that the next measured quantity value cannot be predicted exactly from previous such values. (If a prediction were possible, allowance for the effect could be made!) In general, there can be a number of contributions to each type of error.

3.6 A challenge in measurement is how best to express what is learned about the measurand. Expression of systematic and random error values relating to the measurement, along with a best estimate of the measurand, is one approach that was often used prior to the introduction of the GUM. The GUM provided a different way of thinking about measurement, in particular about how to express the perceived quality of the result of a measurement. Rather than express the result of a measurement by providing a best estimate of the measurand, along with information about systematic and random error values (in the form of an 'error analysis'), the GUM approach is to express the result of a measurement as a best estimate of the measurand, along with an associated *measurement uncertainty*.

3.7 One of the basic premises of the GUM approach is that it is possible to characterize the quality of a measurement by accounting for both systematic and random errors on a comparable footing, and a method is provided for doing that (see 7.2). This method refines the information previously provided in an 'error analysis', and puts it on a probabilistic basis through the concept of measurement uncertainty.

3.8 Another basic premise of the GUM approach is that it is not possible to state how well the essentially unique true value of the measurand is known, but only how well it is believed to be known. Measurement uncertainty can therefore be described as a measure of how well one believes one knows the essentially unique true value of the measurand. This uncertainty reflects the incomplete knowledge of the measurand. The notion of 'belief' is an important one, since it moves metrology into a realm where results of measurement need to be considered and quantified in terms of *probabilities* that express degrees of belief.

3.9 The above discussion concerns the direct measurement of a quantity, which incidentally occurs rarely. The bathroom scales may convert a measured extension of a spring into an estimate of the measurand, the mass of the person on the scales. The particular relationship between extension and mass is determined by the *calibration* [JCGM 200:2008 (VIM) 2.39] of the scales.

3.10 A relationship such as that in 3.9 constitutes a rule for converting a quantity value into the corresponding value of the measurand. The rule is usually known as a *measurement model* [JCGM 200:2008 (VIM) 2.48] or simply a model. There are many types of measurement in practice and therefore many rules or models. Even for one particular type of measurement there may well be more than one model. A simple model (for example a proportional rule, where the mass is proportional to the extension of the spring) might be sufficient for everyday domestic use. Alternatively, a more sophisticated model of a weighing, involving additional effects such as air buoyancy, is capable of delivering better results for industrial or scientific purposes. In general there are often several different quantities, for example temperature, humidity and displacement, that contribute to the definition of the measurand, and that need to be measured.

3.11 Correction terms should be included in the model when the conditions of measurement are not exactly as stipulated. These terms correspond to systematic error values [JCGM 200:2008 (VIM) 2.17]. Given an estimate of a correction term, the relevant quantity should be corrected by this estimate [JCGM 100:2008 (GUM) 3.2.4]. There will be an uncertainty associated with the estimate, even if the estimate is zero, as is often the case. Instances of systematic errors arise in height measurement, when the alignment of the measuring instrument is not perfectly vertical, and the ambient temperature is different from that prescribed. Neither the alignment of the instrument nor the ambient temperature is specified exactly, but information concerning these effects is available, for example the lack of alignment is at most $0.001°$ and the ambient temperature at the time of measurement differs from that stipulated by at most $2\,°C$.

3.12 A quantity can depend on time, for instance a radionuclide decaying at a particular rate. Such an effect should be incorporated into the model to yield a measurand corresponding to a measurement at a given time.

3.13 As well as raw data representing measured quantity values, there is another form of data that is frequently needed in a model. Some such data relate to quantities representing physical constants, each of which is known imperfectly. Examples are material constants such as modulus of elasticity and specific heat. There are often other relevant data given in reference books, calibration certificates, etc., regarded as estimates of further quantities.

3.14 The items required by a model to define a measurand are known as *input quantities in a measurement model* [JCGM 200:2008 (VIM) 2.50]. The rule or model is often referred to as a *functional relationship* [JCGM 100:2008 (GUM) 4.1]. The *output quantity in a measurement model* [JCGM 200:2008 (VIM) 2.51] is the measurand.

3.15 Formally, the output quantity, denoted by Y, about which information is required, is often related to input quantities, denoted by $X_1, \ldots, X_N$, about which information is available, by a measurement model [JCGM 100:2008 (GUM) 4.1.1] in the form of a measurement function [JCGM 200:2008 (VIM) 2.49]

$$Y = f(X_1, \ldots, X_N). \tag{1}$$

3.16 A general expression for a measurement model [JCGM 200:2008 (VIM) 2.48 note 1] is

$$h(Y, X_1, \ldots, X_N) = 0. \tag{2}$$

It is taken that a procedure exists for calculating Y given $X_1, \ldots, X_N$ in equation (2), and that Y is uniquely defined by this equation.

3.17 The true values of the input quantities $X_1, \ldots, X_N$ are unknown. In the approach advocated, $X_1, \ldots, X_N$ are characterized by probability distributions [JCGM 100:2008 (GUM) 3.3.5, ISO 3534-1:2006 2.11] and treated mathematically as random variables [ISO 3534-1:2006 2.10]. These distributions describe the respective probabilities of their true values lying in different intervals, and are assigned based on available knowledge concerning $X_1, \ldots, X_N$. Sometimes, some or all of $X_1, \ldots, X_N$ are interrelated and the relevant distributions, which are known as *joint*, apply to these quantities taken together. The following considerations, which largely apply to unrelated (independent) quantities, can be extended to interrelated quantities.

3.18 Consider estimates $x_1, \ldots, x_N$, respectively, of the input quantities $X_1, \ldots, X_N$, obtained from certificates and reports, manufacturers' specifications, the analysis of measurement data, and so on. The probability distributions characterizing $X_1, \ldots, X_N$ are chosen such that the estimates $x_1, \ldots, x_N$, respectively, are the *expectations* [JCGM 101:2008 3.6, ISO 3534-1:2006 2.12] of $X_1, \ldots, X_N$. Moreover, for the ith input quantity, consider a so-called *standard uncertainty* [JCGM 200:2008 (VIM) 2.30], given the symbol $u(x_i)$, defined as the standard deviation [JCGM 101:2008 3.8, ISO 3534-1:2006 2.37] of the input quantity X_i. This standard uncertainty is said to be *associated* with the (corresponding) estimate x_i. The estimate x_i is best in the sense that $u^2(x_i)$ is smaller than the expected squared difference of X_i from any other value.

3.19 The use of available knowledge to establish a probability distribution to characterize each quantity of interest applies to the X_i and also to Y. In the latter case, the characterizing probability distribution for Y is determined by the functional relationship (1) or (2) together with the probability distributions for the X_i. The determination of the probability distribution for Y from this information is known as the *propagation of distributions* [JCGM 101:2008 5.2].

3.20 Prior knowledge about the true value of the output quantity Y can also be considered. For the domestic bathroom scales, the fact that the person's mass is positive, and that it is the mass of a person, rather than that of a motor car, that is being measured, both constitute prior knowledge about the possible values of the measurand in this example. Such additional information can be used to provide a probability distribution for Y that can give a smaller standard deviation for Y and hence a smaller standard uncertainty associated with the estimate of Y [2, 13, 24].

4 Concepts and basic principles

4.1 Further to those in clause 3, fundamental concepts and principles of probability theory that underlie the approach advocated for the evaluation and expression of measurement uncertainty are provided in JCGM 105:2008 [4].

4.2 Measurement uncertainty is defined [JCGM 200:2008 (VIM) 2.26] as

non-negative parameter characterizing the dispersion of the quantity values being attributed to a measurand, based on the information used.

This definition is consistent with the considerations of 3.8 and 3.17 to 3.20.

4.3 Two representations of a probability distribution [JCGM 101:2008 3.1, ISO 3534-1:2006 2.11] for a random variable X are used in uncertainty evaluation:

— the *distribution function* [JCGM 101:2008 3.2, ISO 3534-1:2006 2.7], a function giving, for every value of its argument, the probability that X be less than or equal to that value, and

— the *probability density function* [JCGM 101:2008 3.3, ISO 3534-1:2006 2.26], the derivative of the distribution function.

4.4 Knowledge of each input quantity X_i in a measurement model is often summarized by the best estimate x_i and the associated standard uncertainty $u(x_i)$ (see 3.18). If, for any i and j, X_i and X_j are related (dependent), the summarizing information will also include a measure of the strength of this relationship, specified as a covariance [ISO 3534-1:2006 2.43] or a correlation. If X_i and X_j are unrelated (independent), their covariance is zero.

4.5 The *evaluation of measurement data*, in the context of the measurement model (1) or (2), is the use of available knowledge concerning the input quantities $X_1, \ldots, X_N$, as represented by the probability distributions used to characterize them, to deduce the corresponding distribution that characterizes the output quantity Y. The evaluation of measurement data might entail determining only a summarizing description of the latter distribution.

4.6 Knowledge about an input quantity X_i is inferred from repeated indication values (*Type A evaluation of uncertainty*) [JCGM 100:2008 (GUM) 4.2, JCGM 200:2008 (VIM) 2.28], or scientific judgement or other information concerning the possible values of the quantity (*Type B evaluation of uncertainty*) [JCGM 100:2008 (GUM) 4.3, JCGM 200:2008 (VIM) 2.29].

4.7 In Type A evaluations of measurement uncertainty [JCGM 200:2008 (VIM) 2.28], the assumption is often made that the distribution best describing an input quantity X given repeated indication values of it (obtained independently) is a Gaussian distribution [ISO 3534-1:2006 2.50]. X then has expectation equal to the average indication value and standard deviation equal to the standard deviation of the average. When the uncertainty is evaluated from a small number of indication values (regarded as instances of an indication quantity characterized by a Gaussian distribution), the corresponding distribution can be taken as a t-distribution [ISO 3534-1:2006 2.53]. Figure 1 shows a Gaussian distribution and (broken curve) a t-distribution with four degrees of freedom. Other considerations apply when the indication values are not obtained independently.

4.8 For a Type B evaluation of uncertainty [JCGM 200:2008 (VIM) 2.29], often the only available information is that X lies in a specified interval $[a, b]$. In such a case, knowledge of the quantity can be characterized by a rectangular probability distribution [JCGM 100:2008 (GUM) 4.3.7, ISO 3534-1:2006 2.60] with limits a and b (figure 2). If different information were available, a probability distribution consistent with that information would be used [26].

4.9 Once the input quantities $X_1, \ldots, X_N$ have been characterized by appropriate probability distributions, and the measurement model has been developed, the probability distribution for the measurand Y is fully specified in terms of this information (also see 3.19). In particular, the expectation of Y is used as the estimate of Y, and the standard deviation of Y as the standard uncertainty associated with this estimate.

4.10 Figure 3 depicts the additive measurement function $Y = X_1 + X_2$ in the case where X_1 and X_2 are each characterized by a (different) rectangular probability distribution. Y has a symmetric trapezoidal probability distribution in this case.

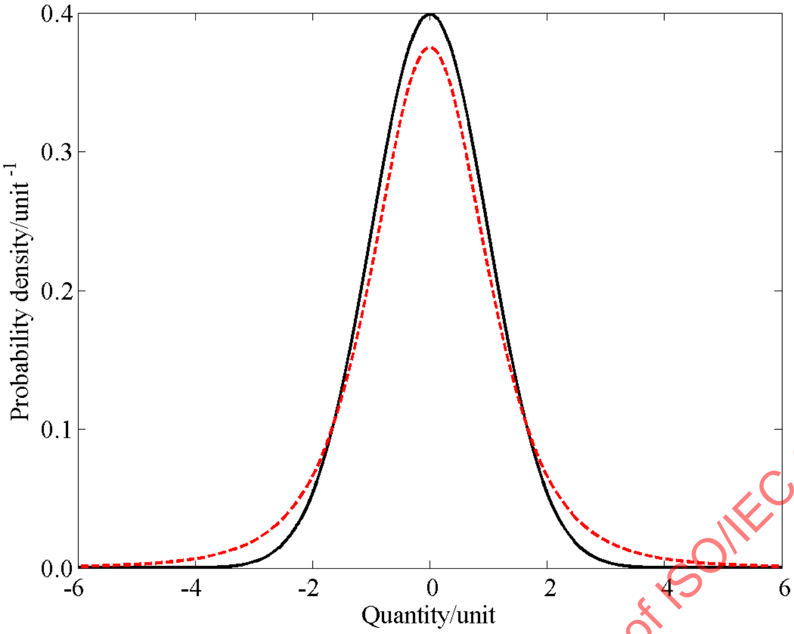

Figure 1 — A Gaussian distribution (continuous black curve) and a t-distribution with four degrees of freedom (broken red curve) (‘unit’ denotes any unit)

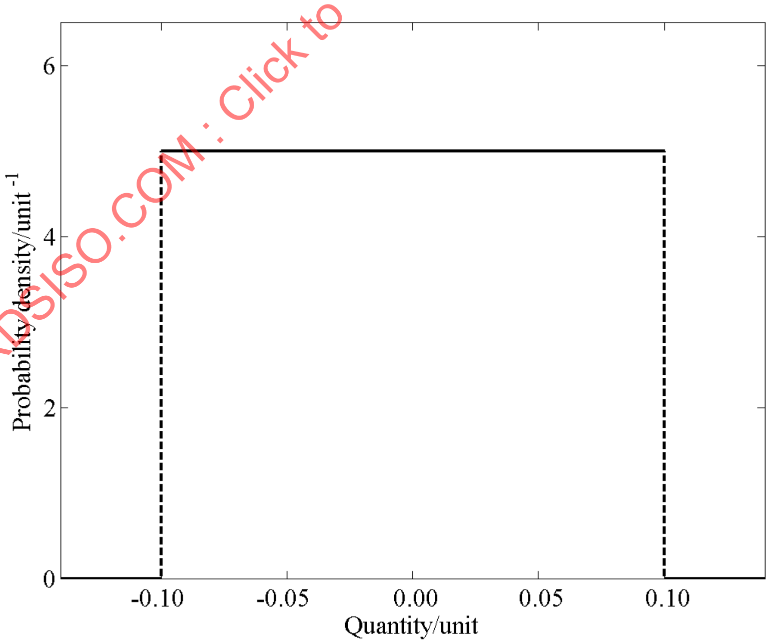

Figure 2 — Rectangular probability distribution with limits -0.1 unit and 0.1 unit (‘unit’ denotes any unit)

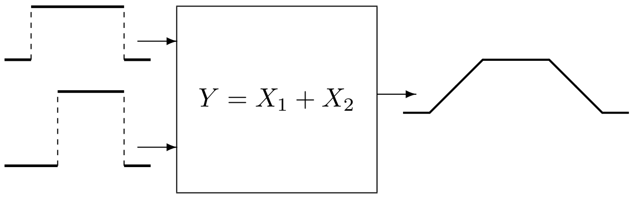

Figure 3 — An additive measurement function with two input quantities X_1 and X_2 characterized by rectangular probability distributions

4.11 Often an interval containing Y with a specified probability is required. Such an interval, a *coverage interval* [JCGM 200:2008 (VIM) 2.36], can be deduced from the probability distribution for Y. The specified probability is known as the *coverage probability* [JCGM 200:2008 (VIM) 2.37].

4.12 For a given coverage probability, there is more than one coverage interval,

a) the *probabilistically symmetric coverage interval* [JCGM 101:2008 3.15], for which the probabilities (summing to one minus the coverage probability) of a value to the left and the right of the interval are equal, and

b) the *shortest coverage interval* [JCGM 101:2008 3.16], for which the length is least over all coverage intervals having the same coverage probability.

4.13 Figure 4 shows a probability distribution (a truncated and scaled Gaussian distribution, indicated by the decreasing curve) with the endpoints of the shortest (continuous blue vertical lines) and those of the probabilistically symmetric (broken red vertical lines) 95 % coverage intervals for a quantity characterized by this distribution. The distribution is asymmetric and the two coverage intervals are different (most notably their right-hand endpoints). The shortest coverage interval has its left-hand endpoint at zero, the smallest possible value for the quantity. The probabilistically symmetric coverage interval in this case is 15 % longer than the shortest coverage interval.

4.14 *Sensitivity coefficients* $c_1, \ldots, c_N$ [JCGM 100:2008 (GUM) 5.1.3] describe how the estimate y of Y would be influenced by small changes in the estimates $x_1, \ldots, x_N$ of the input quantities $X_1, \ldots, X_N$. For the measurement function (1), c_i equals the partial derivative of first order of f with respect to X_i evaluated at $X_1 = x_1, X_2 = x_2$, etc. For the linear measurement function

$$Y = c_1 X_1 + \cdots + c_N X_N, \tag{3}$$

with $X_1, \ldots, X_N$ independent, a change in x_i equal to $u(x_i)$ would give a change $c_i u(x_i)$ in y. This statement would generally be approximate for the measurement models (1) and (2) (see 7.2.4). The relative magnitudes of the terms $|c_i|u(x_i)$ are useful in assessing the respective contributions from the input quantities to the standard uncertainty $u(y)$ associated with y.

4.15 The standard uncertainty $u(y)$ associated with the estimate y of the output quantity Y is not given by the sum of the $|c_i|u(x_i)$, but these terms combined in quadrature [JCGM 100:2008 (GUM) 5.1.3], namely by (an expression that is generally approximate for the measurement models (1) and (2))

$$u^2(y) = c_1^2 u^2(x_1) + \cdots + c_N^2 u^2(x_N). \tag{4}$$

4.16 When the input quantities X_i contain dependencies, formula (4) is augmented by terms containing covariances [JCGM 100:2008 (GUM) 5.2.2], which may increase or decrease $u(y)$.

4.17 According to Resolution 10 of the 22nd CGPM (2003) " . . . the symbol for the decimal marker shall be either the point on the line or the comma on the line . . . ". The JCGM has decided to adopt, in its documents

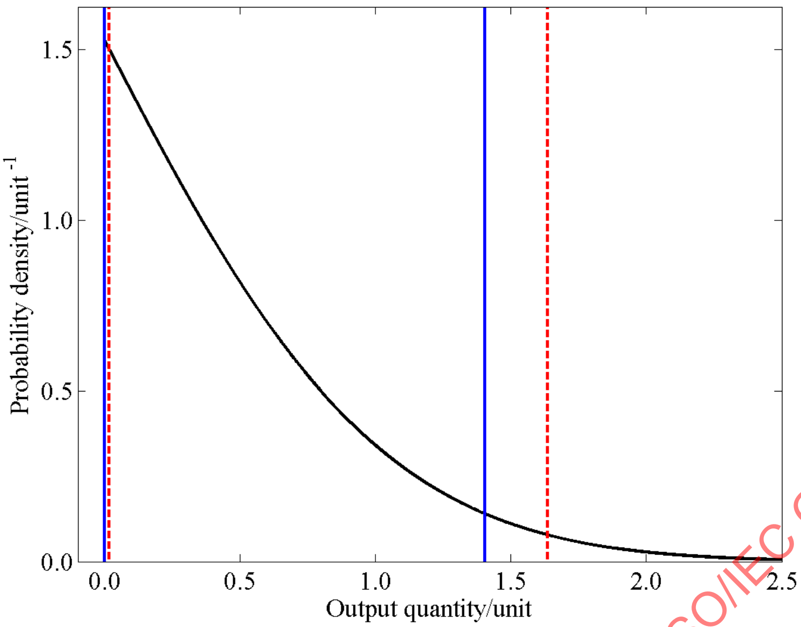

Figure 4 — Shortest 95 % coverage interval (endpoints shown by continuous blue vertical lines) and probabilistically symmetric 95 % coverage interval (broken red) for a quantity characterized by a truncated and scaled Gaussian distribution ('unit' denotes any unit)

in English, the point on the line.

5 Stages of uncertainty evaluation

5.1 The main stages of uncertainty evaluation constitute formulation and calculation, the latter consisting of propagation and summarizing.

5.2 The formulation stage (see clause 6) constitutes

a) defining the output quantity Y (the measurand),

b) identifying the input quantities on which Y depends,

c) developing a measurement model relating Y to the input quantities, and

d) on the basis of available knowledge, assigning probability distributions — Gaussian, rectangular, etc. — to the input quantities (or a joint probability distribution to those input quantities that are not independent).

5.3 The calculation stage (see clause 7) consists of propagating the probability distributions for the input quantities through the measurement model to obtain the probability distribution for the output quantity Y, and summarizing by using this distribution to obtain

a) the expectation of Y, taken as an estimate y of Y,

b) the standard deviation of Y, taken as the standard uncertainty $u(y)$ associated with y [JCGM 100:2008 (GUM) E.3.2], and

c) a coverage interval containing Y with a specified coverage probability.

6 The formulation stage: developing a measurement model

6.1 The formulation stage of uncertainty evaluation involves developing a measurement model, incorporating corrections and other effects as necessary. In some fields of measurement, this stage can be very difficult. It also involves using available knowledge to characterize the input quantities in the model by probability distributions. JCGM 103 [6] provides guidance on developing and working with a measurement model. The assignment of probability distributions to the input quantities in a measurement model is considered in JCGM 101 [JCGM 101:2008 6] and JCGM 102 [5].

6.2 A measurement model relating the input quantities to the output quantity is initially developed. There might be more than one output quantity (see 6.5). The model is formed on theoretical or empirical grounds or both, and generally depends on the metrology discipline, electrical, dimensional, thermal, mass, etc. The model is then augmented by terms constituting further input quantities, describing effects that influence the measurement. JCGM 103 [6] provides guidance on handling these additional effects, which may be categorized into random and systematic effects.

6.3 JCGM 103 considers broader classes of measurement model than does the GUM, categorizing the model according to whether

a) the quantities involved are real or complex,

b) the measurement model takes the general form (2) or can be expressed as a measurement function (1), and

c) there is a single output quantity or more than one output quantity (see 6.5).

In category (a), complex quantities occur especially in electrical metrology, and also in acoustical and optical metrology. In category (b), for a measurement function the output quantity is expressed directly as a formula involving the input quantities, and for a general measurement model an equation is solved for the output quantity in terms of the input quantities (see 6.5).

6.4 Examples from a range of metrology disciplines illustrate various aspects of JCGM 103. Guidance on numerical analysis aspects that arise in treating these examples is given. Guidance also includes the use of changes of variables so that all or some of the resulting quantities are uncorrelated or only weakly correlated.

6.5 The GUM and JCGM 101:2008 concentrate on measurement models in the form of measurement functions having a single output quantity Y. Many measurement problems arise, however, for which there is more than one output quantity, depending on a common set of input quantities. These output quantities are denoted by $Y_1, \ldots, Y_m$. Instances include (a) an output quantity that is complex, and represented in terms of its real and imaginary components (or magnitude and phase), (b) quantities representing the parameters of a calibration function, and (c) quantities describing the geometry of the surface of an artefact. The GUM does not directly address such models, although examples are given concerning simultaneous resistance and reactance measurement [JCGM 100:2008 (GUM) H.2] and thermometer calibration [JCGM 100:2008 (GUM) H.3].

6.6 The formulation phase of uncertainty evaluation for the case of more than one measurand is consistent with that for a measurement model with a single measurand: it comprises developing a model and assigning probability distributions to the input quantities based on available knowledge. As for a measurement model with a single output quantity, there is an estimate of each input quantity and a standard uncertainty associated with that estimate (and possibly covariances associated with pairs of estimates). Furthermore, since in general each output quantity depends on all the input quantities, in addition to determining estimates of these output quantities and the standard uncertainties associated with these estimates, it is required to evaluate the covariances associated with all pairs of these estimates.

6.7 The counterpart of the measurement function (1) for a number m of output quantities is

$$Y_1 = f_1(X_1, \ldots, X_N), \quad Y_2 = f_2(X_1, \ldots, X_N), \; \ldots, \quad Y_m = f_m(X_1, \ldots, X_N), \tag{5}$$

in which there are m functions $f_1, \ldots, f_m$. Figure 5 illustrates such a measurement function.

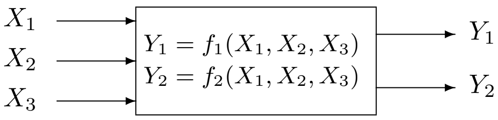

Figure 5 — A measurement function with three input quantities X_1, X_2 and X_3, and two output quantities Y_1 and Y_2

6.8 *Multistage measurement models*, where the output quantities from previous stages become the input quantities to subsequent stages, are also treated in JCGM 103. A common example of a multistage measurement model relates to the construction and use of a calibration function [JCGM 200:2008 (VIM) 2.39] (see figure 6):

a) Given quantity values provided by measurement standards, and corresponding indication values provided by a measuring system, determine estimates of the parameters of the calibration function. The standard uncertainties associated with the measured quantity values and the indication values give rise to standard uncertainties associated with these estimates and in general with covariances associated with all pairs of these estimates;

b) Given a further indication value, evaluate the calibration function to provide the corresponding measured quantity value. This step involves the inverse of the calibration function. The standard uncertainties and covariances associated with the estimates of the parameters of the calibration function, together with the standard uncertainty associated with the further indication value, give rise to a standard uncertainty associated with this measured quantity value.

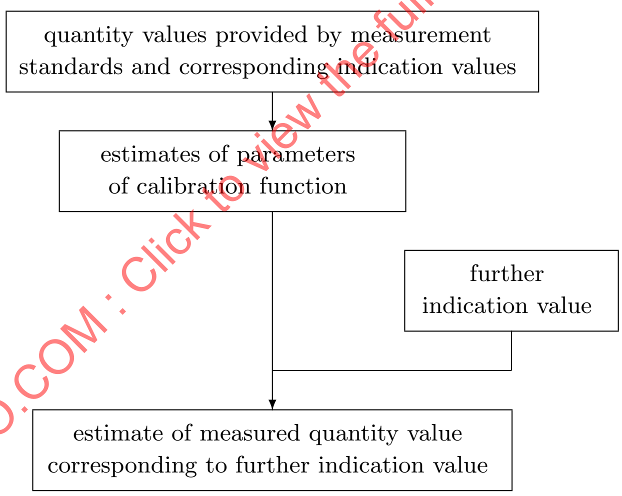

Figure 6 — A two-stage measurement model for a calibration function in which quantity values provided by measurement standards and corresponding indication values are used to establish estimates of the parameters of the calibration function, which, given a further indication value, are used to estimate the corresponding measured quantity value

7 The calculation (propagation and summarizing) stage of uncertainty evaluation

7.1 General

7.1.1 The propagation stage of uncertainty evaluation is known as the *propagation of distributions* [JCGM 101:2008 5.2], various approaches for which are available, including

a) the GUM uncertainty framework, constituting the application of the law of propagation of uncertainty, and the characterization of the output quantity Y by a Gaussian or a t-distribution (see 7.2),

b) analytic methods, in which mathematical analysis is used to derive an algebraic form for the probability distribution for Y (see 7.3), and

c) a Monte Carlo method (MCM), in which an approximation to the distribution function for Y is established numerically by making random draws from the probability distributions for the input quantities, and evaluating the model at the resulting values (see 7.4).

7.1.2 For any particular uncertainty evaluation problem, approach a), b) or c) (or some other approach) is used, a) being generally approximate, b) exact, and c) providing a solution with a numerical accuracy that can be controlled.

7.1.3 The application of approaches a) and c) to measurement functions with any number of output quantities, and general measurement models, is considered in 7.5.

7.2 The GUM uncertainty framework

7.2.1 The GUM uncertainty framework [JCGM 100:2008 (GUM) 3.4.8, 5.1] (depicted in figure 7) uses

a) the best estimates x_i of the input quantities X_i,

b) the standard uncertainties $u(x_i)$ associated with the x_i, and

c) the sensitivity coefficients c_i (see 4.14)

to form an estimate y of the output quantity Y and the associated standard uncertainty $u(y)$.

7.2.2 A variant [JCGM 100:2008 (GUM) 5.2] of 7.2.1 applies when the input quantities are mutually dependent (not indicated in figure 7). By regarding the probability distribution for Y as Gaussian, a coverage interval for Y corresponding to a specified coverage probability is also determined [JCGM 100:2008 (GUM) G.2]. When the degrees of freedom [ISO 3534-1:2006 2.54] relating to any $u(x_i)$ is finite, an (effective) degrees of freedom relating to $u(y)$ is determined, and the probability distribution for Y taken as a t-distribution.

7.2.3 There are many circumstances where the GUM uncertainty framework [JCGM 100:2008 (GUM) 5] can be applied and leads to valid statements of uncertainty. If the measurement function is linear in the input quantities and the probability distributions for these quantities are Gaussian, the GUM uncertainty framework provides exact results [JCGM 101:2008 5.7]. Even when these conditions do not hold, the approach can often work sufficiently well for practical purposes [JCGM 101:2008 5.8].

7.2.4 There are situations where the GUM uncertainty framework might not be satisfactory, including those where

a) the measurement function is non-linear,

b) the probability distributions for the input quantities are asymmetric,

c) the uncertainty contributions $|c_1|u(x_1), \ldots, |c_N|u(x_N)$ (see 4.14) are not of approximately the same magnitude [JCGM 100:2008 (GUM) G.2.2], and

d) the probability distribution for the output quantity is either asymmetric, or not a Gaussian or a t-distribution.

Sometimes it is hard to establish in advance that the circumstances hold for the GUM uncertainty framework to apply.

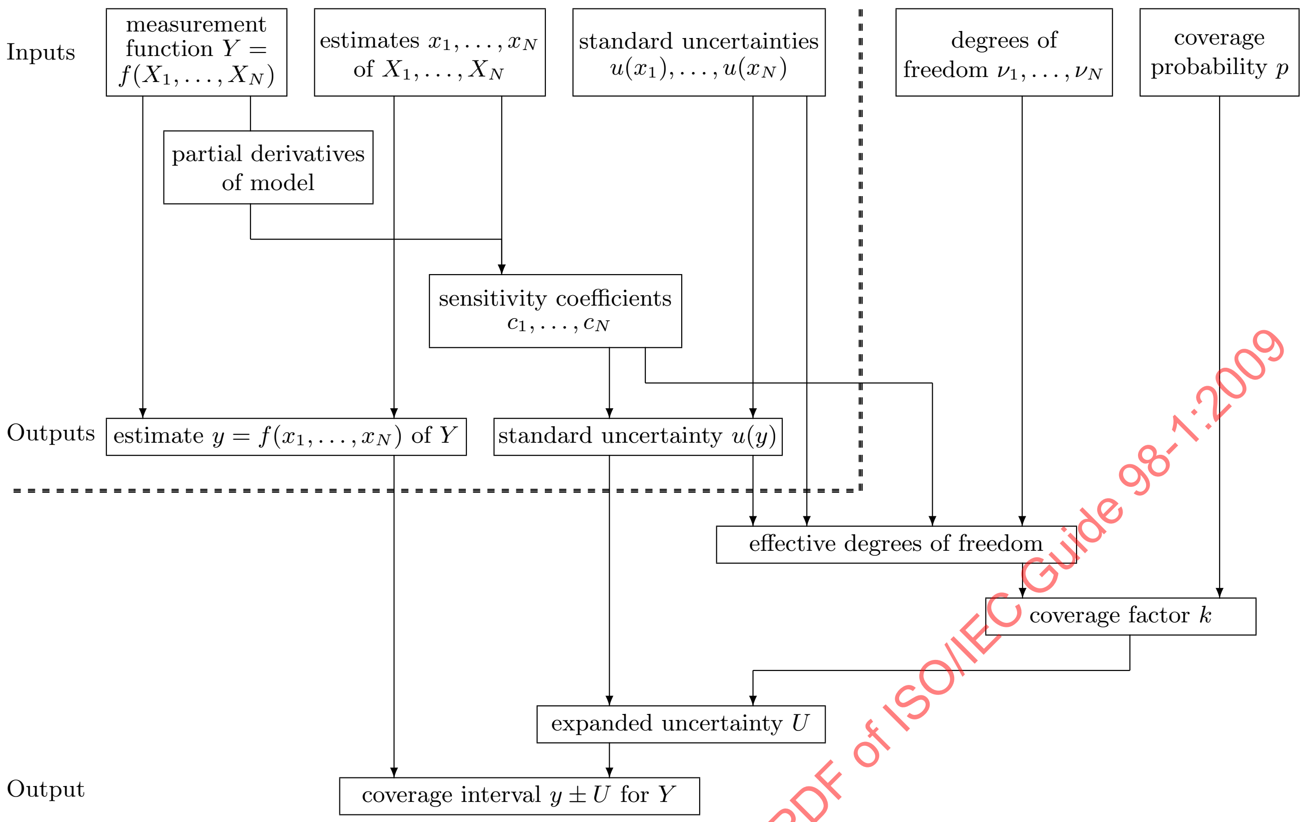

Figure 7 — Measurement uncertainty evaluation using the GUM uncertainty framework, where the top-left part of the figure (bounded by broken lines) relates to obtaining an estimate y of the output quantity Y and the associated standard uncertainty $u(y)$, and the remainder relates to the determination of a coverage interval for Y

7.2.5 The use of the GUM uncertainty framework becomes more difficult when forming partial derivatives (or numerical approximations to them) for a measurement model that is complicated, as needed by the law of propagation of uncertainty (possibly with higher-order terms) [JCGM 100:2008 (GUM) 5]. A valid and sometimes more readily applicable treatment is obtained by applying a suitable Monte Carlo implementation of the propagation of distributions (see 7.4).

7.3 Analytic methods

7.3.1 Analytic methods by which an algebraic form for the probability distribution for the output quantity can be obtained do not introduce any approximation, but can be applied only in relatively simple cases. A treatment of such methods is available [8, 12]. Some cases that can be so handled for a general number N of input quantities are linear measurement functions (expression (3)), where the probability distributions for all input quantities are Gaussian, or all are rectangular with the same width. An instance with two input quantities ($N = 2$), for which the probability distributions for the input quantities are rectangular, and the probability distribution for the output quantity is trapezoidal [10], is illustrated in figure 3.

7.3.2 Cases where there is one input quantity ($N = 1$) can often be treated analytically, using a formula [25, pages 57–61] to derive algebraically a probability distribution for the output quantity. Such cases arise in the transformation of measurement units, for example from linear to logarithmic units [10, pages 95–98].

7.3.3 An advantage of an algebraic solution is that it provides insight through displaying the dependence of the probability distribution for the output quantity on parameters of the probability distributions for the input quantities.

7.4 Monte Carlo method

7.4.1 JCGM 101:2008 provides detailed information on MCM as an implementation of the propagation of distributions [JCGM 101:2008 5.9]. MCM has fewer conditions associated with its use than the GUM uncertainty framework [JCGM 101:2008 5.10]. Figure 8 illustrates the procedure. JCGM 101:2008 gives examples to compare MCM with the use of the GUM uncertainty framework [JCGM 101:2008 9].

7.4.2 JCGM 101:2008 provides an adaptive MCM procedure, in which the number of Monte Carlo trials is determined automatically by utilizing a measure of convergence of the overall process [JCGM 101:2008 7.9].

7.4.3 In JCGM 101:2008 there is a procedure that uses MCM to decide whether the application of the GUM uncertainty framework in any particular case is valid [JCGM 101:2008 8].

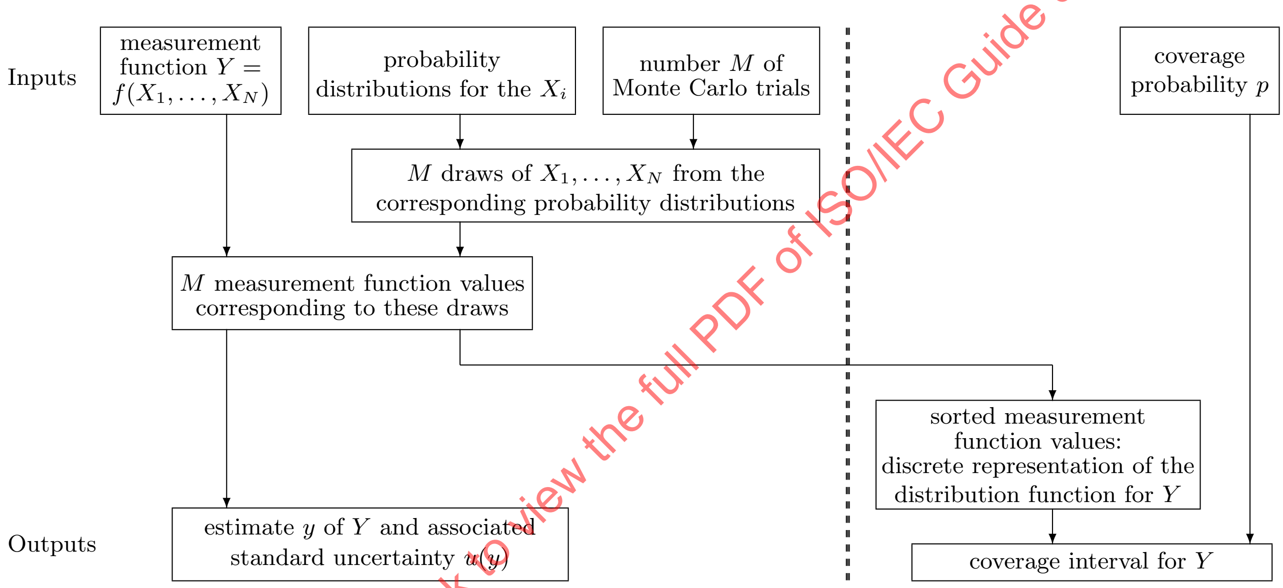

Figure 8 — Measurement uncertainty evaluation using a Monte Carlo method, where the part of the figure to the left of the broken line relates to obtaining an estimate y of the output quantity Y and the associated standard uncertainty $u(y)$, and the remainder relates to the determination of a coverage interval for Y

7.5 Measurement models with any number of output quantities

7.5.1 In order to evaluate the uncertainties and covariances associated with estimates of the output quantities for measurement models with any number of output quantities, both the GUM uncertainty framework and MCM, as treated in JCGM 101:2008, require extension. The GUM [JCGM 100:2008 (GUM) F.1.2.3] outlines such an extension of the GUM uncertainty framework, but considers it further only in examples.

7.5.2 In JCGM 102 [5], it is stated that the law of propagation of uncertainty, a main constituent of the GUM uncertainty framework, can succinctly be expressed in an equivalent matrix form when applied to a measurement model having a single output quantity. The matrix expression has the advantage of being suitable as a basis for implementation in software, and for extension to more general types of measurement model.

7.5.3 That extension is given in JCGM 102 for a measurement function having any number of output quantities. The extension to any number of output quantities in a general measurement model (see 3.16) is also treated in JCGM 102.

7.5.4 JCGM 102 also applies MCM to measurement models with any number of output quantities. A discrete representation of the probability distribution for the output quantities is provided. Expressions are given for the estimates of the output quantities, the standard uncertainties associated with these estimates, and the covariances associated with pairs of these estimates in terms of that representation.

7.5.5 In addition to obtaining estimates of the output quantities, together with the associated standard uncertainties and covariances, it might be required to obtain a *region* containing the output quantities with a specified (coverage) probability. It is natural to consider the extension to regions of the probabilistically symmetric coverage interval and the shortest coverage interval. However, there is no natural counterpart of a probabilistically symmetric coverage interval in the form of a coverage region, whereas there is for a shortest coverage interval. The determination of a smallest coverage region is generally a difficult task.

7.5.6 In some circumstances, it is reasonable to provide an *approximate* coverage region having simple geometric shape. Two particular forms of coverage region are considered in this regard. One form results from characterizing the output quantities by a joint Gaussian distribution, for example on the basis of the central limit theorem [JCGM 100:2008 (GUM) G.2], in which case the smallest coverage region is bounded by a hyper-ellipsoid. The other form constitutes a hyper-rectangular coverage region. Procedures for obtaining these forms are provided in JCGM 102.

8 Measurement uncertainty in conformity assessment

8.1 Conformity assessment is an area of importance in manufacturing quality control, legal metrology, and in the maintenance of health and safety. In the industrial inspection of manufactured parts, decisions are made concerning the compatibility of the parts with the design specification. Similar issues arise in the context of regulation (relating to emissions, radiation, drugs, doping control, etc.) concerning whether stipulated limits for true quantity values have been surpassed. Guidance is provided in JCGM 106 [7]. Also see reference [18].

8.2 Measurement is intrinsic to conformity assessment in deciding whether the output quantity, or measurand, conforms to a specified requirement. For a single quantity, such a requirement typically takes the form of specification limits that define an interval of permissible quantity values. In the absence of uncertainty, a measured quantity value lying within this interval is said to be conforming, and non-conforming otherwise. The influence of measurement uncertainty on the inspection process necessitates a balance of risks between producers and consumers.

8.3 The possible values of a quantity Y of interest are represented by a probability distribution. The probability that Y conforms to specification can be calculated, given this probability distribution and the specification limits.

8.4 Because of the incomplete knowledge of the quantity Y (as encoded in its probability distribution), there is a risk of a mistaken decision in deciding conformity to specification. Such mistaken decisions are of two types: a quantity accepted as conforming might actually be non-conforming, and a quantity rejected as non-conforming might actually conform. The related risks correspond, respectively, to *consumer's risk* and *producer's risk* (see JCGM 106).

8.5 By defining an *acceptance interval* of acceptable measured quantity values, the risks of a mistaken decision concerning acceptance or rejection can be balanced so as to minimize the costs associated with these decisions [19]. The problem of calculating the conformity probability and the probabilities of the two types of mistaken decision, given the probability distribution, the specification limits, and the limits of the acceptance interval is addressed in JCGM 106. The choice of acceptance interval limits is a matter that depends on the implications of these mistaken decisions.

8.6 Although the probability distribution in 8.3 to 8.5 is general, the treatment is then specialized in JCGM 106 to the most important case in practice, namely, when the probability distribution is Gaussian.

9 Applications of the least-squares method

9.1 Guidance on the application of the least-squares method (also known as least-squares adjustment) to data evaluation problems in metrology is provided in JCGM 107 [3]. In such problems there is often an underlying theoretical relationship between an independent variable and a dependent variable. This relationship constitutes the basis of a parameter adjustment or curve-fitting problem. The input quantities in the related measurement model are the quantities of which the measured values of the independent and dependent variables are outcomes. The output quantities are the quantities representing the required parameters. The manner in which the output quantities are obtained from the input quantities by means of a least-squares procedure defines the measurement model.

9.2 In calibration terminology (see 6.8), a measured quantity value of an independent variable would typically be that of a measurement standard. The value of the dependent variable would be an indication value returned by the measuring system for the corresponding value of the independent variable. In the curve-fitting context, which includes calibration as a special case, the adjustment procedure used in JCGM 107 is a generalized version of the usual least-squares procedure.

9.3 The task is to estimate the parameters (and sometimes even their number) from pairs of measured quantity values and the corresponding indication values. These pairs, together with the associated standard uncertainties and, when appropriate, covariances, constitute the input data to the adjustment.

9.4 Typical measurement problems to which JCGM 107 can be applied include (a) linear or non-linear curve-fitting problems, including the case of imperfectly known values of the independent variable, and (b) fitting of general models to estimate parameters in a physical process. The application of JCGM 107 is not restricted to curve-fitting problems in the strictest sense. It can also be used to treat, for instance, convolution problems [21], the adjustment of fundamental constants [22], and key comparison data evaluation [9].

9.5 For problems of type (a) in 9.4, once the least-squares method has been used to estimate the parameters of a calibration function and evaluate the associated standard uncertainties and covariances, the measuring system will subsequently be used for measurement. The estimates of the parameters of the calibration function, together with a particular indication value, are then used to estimate the corresponding quantity. The standard uncertainty associated with this estimate is evaluated using the standard uncertainties and covariances associated with the parameter estimates and the standard uncertainty associated with the indication value.

9.6 It is emphasized in JCGM 107 that the *uncertainty structure* should be taken fully into account when formulating and solving the least-squares problem. 'Uncertainty structure' refers to the standard uncertainties associated with the measured quantity values and indication values and any covariances associated with pairs of these values.

9.7 For problems of type (b) in 9.4, or in terms of determining the parameters in problems of type (a), the adjustment problem is rarely a problem in only one output quantity. Rather, the problem involves a number of output quantities in which the mathematical formulation can conveniently be expressed in terms of matrices. JCGM 107 makes extensive use of matrix formalism, which is well adapted to numerical solution using a computer, as usually required in practice (also see 7.5).