

# GUIDE TO THE EXPRESSION OF UNCERTAINTY IN MEASUREMENT

CORRECTED AND REPRINTED, 1995

BIPM International Bureau of Weights and Measures  
IEC International Electrotechnical Commission  
IFCC International Federation of Clinical Chemistry  
ISO International Organization for Standardization  
IUPAC International Union of Pure and Applied Chemistry  
IUPAP International Union of Pure and Applied Physics  
OIML International Organization of Legal Metrology

- BIPM Bureau international des poids et mesures  
Pavillon de Breteuil  
F-92312 Sèvres Cedex  
France
- IEC International Electrotechnical Commission  
3, rue de Varembe  
Case postale 131  
CH-1211 Genève 20  
Switzerland
- IFCC International Federation of Clinical Chemistry  
Technical Secretariat  
Centre du Médicament  
Université de Nancy 1  
30, rue Lionnois  
F-54000 Nancy  
France
- ISO International Organization for Standardization  
1, rue de Varembe  
Case postale 56  
CH-1211 Genève 20  
Switzerland
- IUPAC International Union of Pure and Applied Chemistry  
Bank Court Chambers  
2-3 Pound Way  
Templars Square, Cowley  
Oxford OX4 3YF  
United Kingdom
- IUPAP International Union of Pure and Applied Physics  
Secretariat  
Vittens gata 11  
S-421 65 V. Frölunda  
Sweden
- OIML International Organization of Legal Metrology  
11, rue Turgot  
F-75009 Paris  
France

# GUIDE TO THE EXPRESSION OF UNCERTAINTY IN MEASUREMENT

This *Guide* establishes general rules for evaluating and expressing uncertainty in measurement that are intended to be applicable to a broad spectrum of measurements. The basis of the *Guide* is Recommendation 1 (CI-1981) of the Comité International des Poids et Mesures (CIPM) and Recommendation INC-1 (1980) of the Working Group on the Statement of Uncertainties. The Working Group was convened by the Bureau International des Poids et Mesures (BIPM) in response to a request of the CIPM. The CIPM Recommendation is the only recommendation concerning the expression of uncertainty in measurement adopted by an intergovernmental organization.

This *Guide* was prepared by a joint working group consisting of experts nominated by the BIPM, the International Electrotechnical Commission (IEC), the International Organization for Standardization (ISO), and the International Organization of Legal Metrology (OIML).

The following seven organizations supported the development of this *Guide*, which is published in their name:

|       |                                                   |
|-------|---------------------------------------------------|
| BIPM  | Bureau International des Poids et Mesures         |
| IEC   | International Electrotechnical Commission         |
| IFCC  | International Federation of Clinical Chemistry    |
| ISO   | International Organization for Standardization    |
| IUPAC | International Union of Pure and Applied Chemistry |
| IUPAP | International Union of Pure and Applied Physics   |
| OIML  | International Organization of Legal Metrology     |

Users of this *Guide* are invited to send their comments and requests for clarification to any of the seven supporting organizations, the mailing addresses of which are given on the inside front cover.

Guide to the Expression of Uncertainty in Measurement

First edition 1995  
ISBN 92-67-10188-9

© International Organization for Standardization  
1995

Printed in Switzerland

# Contents

|                                                                            | Page |                                                                                        | Page |
|----------------------------------------------------------------------------|------|----------------------------------------------------------------------------------------|------|
| Foreword . . . . .                                                         | v    |                                                                                        |      |
| 0 Introduction . . . . .                                                   | vii  |                                                                                        |      |
| 1 Scope . . . . .                                                          | 1    |                                                                                        |      |
| 2 Definitions . . . . .                                                    | 2    | <b>Annexes</b>                                                                         |      |
| 2.1 General metrological terms . . . . .                                   | 2    | <b>A Recommendations of Working Group and CIPM . . . . .</b>                           | 29   |
| 2.2 The term “uncertainty” . . . . .                                       | 2    | <b>A.1 Recommendation INC-1 (1980) . . . . .</b>                                       | 29   |
| 2.3 Terms specific to this <i>Guide</i> . . . . .                          | 3    | <b>A.2 Recommendation 1 (CI-1981) . . . . .</b>                                        | 30   |
| 3 Basic concepts . . . . .                                                 | 4    | <b>A.3 Recommendation 1 (CI-1986) . . . . .</b>                                        | 30   |
| 3.1 Measurement . . . . .                                                  | 4    | <b>B General metrological terms . . . . .</b>                                          | 31   |
| 3.2 Errors, effects, and corrections . . . . .                             | 5    | <b>B.1 Source of definitions . . . . .</b>                                             | 31   |
| 3.3 Uncertainty . . . . .                                                  | 5    | <b>B.2 Definitions . . . . .</b>                                                       | 31   |
| 3.4 Practical considerations . . . . .                                     | 7    | <b>C Basic statistical terms and concepts . . . . .</b>                                | 35   |
| 4 Evaluating standard uncertainty . . . . .                                | 9    | <b>C.1 Source of definitions . . . . .</b>                                             | 35   |
| 4.1 Modelling the measurement . . . . .                                    | 9    | <b>C.2 Definitions . . . . .</b>                                                       | 35   |
| 4.2 Type A evaluation of standard uncertainty . . . . .                    | 10   | <b>C.3 Elaboration of terms and concepts . . . . .</b>                                 | 38   |
| 4.3 Type B evaluation of standard uncertainty . . . . .                    | 11   | <b>D “True” value, error, and uncertainty . . . . .</b>                                | 40   |
| 4.4 Graphical illustration of evaluating standard uncertainty . . . . .    | 14   | <b>D.1 The measurand . . . . .</b>                                                     | 40   |
| 5 Determining combined standard uncertainty . . . . .                      | 19   | <b>D.2 The realized quantity . . . . .</b>                                             | 40   |
| 5.1 Uncorrelated input quantities . . . . .                                | 19   | <b>D.3 The “true” value and the corrected value . . . . .</b>                          | 40   |
| 5.2 Correlated input quantities . . . . .                                  | 20   | <b>D.4 Error . . . . .</b>                                                             | 41   |
| 6 Determining expanded uncertainty . . . . .                               | 23   | <b>D.5 Uncertainty . . . . .</b>                                                       | 41   |
| 6.1 Introduction . . . . .                                                 | 23   | <b>D.6 Graphical representation . . . . .</b>                                          | 42   |
| 6.2 Expanded uncertainty . . . . .                                         | 23   | <b>E Motivation and basis for Recommendation INC-1 (1980) . . . . .</b>                | 45   |
| 6.3 Choosing a coverage factor . . . . .                                   | 24   | <b>E.1 “Safe,” “random,” and “systematic” . . . . .</b>                                | 45   |
| 7 Reporting uncertainty . . . . .                                          | 25   | <b>E.2 Justification for realistic uncertainty evaluations . . . . .</b>               | 45   |
| 7.1 General guidance . . . . .                                             | 25   | <b>E.3 Justification for treating all uncertainty components identically . . . . .</b> | 46   |
| 7.2 Specific guidance . . . . .                                            | 25   | <b>E.4 Standard deviations as measures of uncertainty . . . . .</b>                    | 48   |
| 8 Summary of procedure for evaluating and expressing uncertainty . . . . . | 28   | <b>E.5 A comparison of two views of uncertainty . . . . .</b>                          | 49   |

|                                                                                                                 |    |                                                                        |    |
|-----------------------------------------------------------------------------------------------------------------|----|------------------------------------------------------------------------|----|
| <b>F</b> Practical guidance on evaluating uncertainty components . . . . .                                      | 51 | <b>H</b> Examples . . . . .                                            | 67 |
| <b>F.1</b> Components evaluated from repeated observations: Type A evaluation of standard uncertainty . . . . . | 51 | <b>H.1</b> End-gauge calibration . . . . .                             | 67 |
| <b>F.2</b> Components evaluated by other means: Type B evaluation of standard uncertainty . . . . .             | 53 | <b>H.2</b> Simultaneous resistance and reactance measurement . . . . . | 72 |
| <b>G</b> Degrees of freedom and levels of confidence . . . . .                                                  | 59 | <b>H.3</b> Calibration of a thermometer . . . . .                      | 76 |
| <b>G.1</b> Introduction . . . . .                                                                               | 59 | <b>H.4</b> Measurement of activity . . . . .                           | 79 |
| <b>G.2</b> Central Limit Theorem . . . . .                                                                      | 60 | <b>H.5</b> Analysis of variance . . . . .                              | 83 |
| <b>G.3</b> The <i>t</i> -distribution and degrees of freedom . . . . .                                          | 60 | <b>H.6</b> Measurements on a reference scale: hardness . . . . .       | 88 |
| <b>G.4</b> Effective degrees of freedom . . . . .                                                               | 61 | <b>J</b> Glossary of principal symbols . . . . .                       | 91 |
| <b>G.5</b> Other considerations . . . . .                                                                       | 63 | <b>K</b> Bibliography . . . . .                                        | 94 |
| <b>G.6</b> Summary and conclusions . . . . .                                                                    | 64 | <b>Alphabetical index</b> . . . . .                                    | 96 |

## Foreword

In 1977, recognizing the lack of international consensus on the expression of uncertainty in measurement, the world's highest authority in metrology, the Comité International des Poids et Mesures (CIPM), requested the Bureau International des Poids et Mesures (BIPM) to address the problem in conjunction with the national standards laboratories and to make a recommendation.

The BIPM prepared a detailed questionnaire covering the issues involved and distributed it to 32 national metrology laboratories known to have an interest in the subject (and, for information, to five international organizations). By early 1979 responses were received from 21 laboratories [1].<sup>1</sup> Almost all believed that it was important to arrive at an internationally accepted procedure for expressing measurement uncertainty and for combining individual uncertainty components into a single total uncertainty. However, a consensus was not apparent on the method to be used. The BIPM then convened a meeting for the purpose of arriving at a uniform and generally acceptable procedure for the specification of uncertainty; it was attended by experts from 11 national standards laboratories. This Working Group on the Statement of Uncertainties developed Recommendation INC-1 (1980), Expression of Experimental Uncertainties [2]. The CIPM approved the Recommendation in 1981 [3] and reaffirmed it in 1986 [4].

The task of developing a detailed guide based on the Working Group Recommendation (which is a brief outline rather than a detailed prescription) was referred by the CIPM to the International Organization for Standardization (ISO), since ISO could better reflect the needs arising from the broad interests of industry and commerce.

Responsibility was assigned to the ISO Technical Advisory Group on Metrology (TAG 4) because one of its tasks is to coordinate the development of guidelines on measurement topics that are of common interest to ISO and the six organizations that participate with ISO in the work of TAG 4: the International Electrotechnical Commission (IEC), the partner of ISO in worldwide standardization; the CIPM and the International Organization of Legal Metrology (OIML), the two worldwide metrology organizations; the International Union of Pure and Applied Chemistry (IUPAC) and the International Union of Pure and Applied Physics (IUPAP), the two international unions that represent chemistry and physics; and the International Federation of Clinical Chemistry (IFCC).

TAG 4 in turn established Working Group 3 (ISO/TAG 4/WG 3) composed of experts nominated by the BIPM, IEC, ISO, and OIML and appointed by the Chairman of TAG 4. It was assigned the following terms of reference:

To develop a guidance document based upon the recommendation of the BIPM Working Group on the Statement of Uncertainties which provides rules on the expression of measurement uncertainty for use within standardization, calibration, laboratory accreditation, and metrology services;

The purpose of such guidance is

- to promote full information on how uncertainty statements are arrived at;
- to provide a basis for the international comparison of measurement results.

1) See the bibliography on page 94 *et seq.*



## 0 Introduction

**0.1** When reporting the result of a measurement of a physical quantity, it is obligatory that some quantitative indication of the quality of the result be given so that those who use it can assess its reliability. Without such an indication, measurement results cannot be compared, either among themselves or with reference values given in a specification or standard. It is therefore necessary that there be a readily implemented, easily understood, and generally accepted procedure for characterizing the quality of a result of a measurement, that is, for evaluating and expressing its *uncertainty*.

**0.2** The concept of *uncertainty* as a quantifiable attribute is relatively new in the history of measurement, although *error* and *error analysis* have long been a part of the practice of measurement science or metrology. It is now widely recognized that, when all of the known or suspected components of error have been evaluated and the appropriate corrections have been applied, there still remains an uncertainty about the correctness of the stated result, that is, a doubt about how well the result of the measurement represents the value of the quantity being measured.

**0.3** Just as the nearly universal use of the International System of Units (SI) has brought coherence to all scientific and technological measurements, a worldwide consensus on the evaluation and expression of uncertainty in measurement would permit the significance of a vast spectrum of measurement results in science, engineering, commerce, industry, and regulation to be readily understood and properly interpreted. In this era of the global marketplace, it is imperative that the method for

evaluating and expressing uncertainty be uniform throughout the world so that measurements performed in different countries can be easily compared.

**0.4** The ideal method for evaluating and expressing the uncertainty of the result of a measurement should be:

- *universal*: the method should be applicable to all kinds of measurements and to all types of input data used in measurements.

The actual quantity used to express uncertainty should be:

- *internally consistent*: it should be directly derivable from the components that contribute to it, as well as independent of how these components are grouped and of the decomposition of the components into subcomponents;
- *transferable*: it should be possible to use directly the uncertainty evaluated for one result as a component in evaluating the uncertainty of another measurement in which the first result is used.

Further, in many industrial and commercial applications, as well as in the areas of health and safety, it is often necessary to provide an interval about the measurement result that may be expected to encompass a large fraction of the distribution of values that could reasonably be attributed to the quantity subject to measurement. Thus the ideal method for evaluating and expressing uncertainty in measurement should be capable of readily providing such an interval, in particular, one with a coverage probability or level of confidence that corresponds in a realistic way with that required.

**0.5** The approach upon which this guidance document is based is that outlined in Recommendation INC-1 (1980) [2] of the Working Group on the Statement of Uncertainties, which was convened by the BIPM in response to a request of the CIPM (see Foreword). This approach, the justification of which is discussed in annex E, meets all of the requirements outlined above. This is not the case for most other methods in current use. Recommendation INC-1 (1980) was approved and reaffirmed by the CIPM in its own Recommendations 1 (CI-1981) [3] and 1 (CI-1986) [4]; the English translations of these CIPM Recommendations are reproduced in annex A (see A.2 and A.3, respectively). Because Recommendation INC-1 (1980) is the foundation upon which this document rests, the English translation is reproduced in 0.7 and the French text, which is authoritative, is reproduced in A.1.

**0.6** A succinct summary of the procedure specified in this guidance document for evaluating and expressing uncertainty in measurement is given in clause 8 and a number of examples are presented in detail in annex H. Other annexes deal with general terms in metrology (annex B); basic statistical terms and concepts (annex C); “true” value, error, and uncertainty (annex D); practical suggestions for evaluating uncertainty components (annex F); degrees of freedom and levels of confidence (annex G); the principal mathematical symbols used throughout the document (annex J); and bibliographical references (annex K). An alphabetical index concludes the document.

## 0.7 Recommendation INC-1 (1980)

### Expression of experimental uncertainties

1. The uncertainty in the result of a measurement generally consists of several components which may be grouped into two categories according to the way in which their numerical value is estimated:

- A. those which are evaluated by statistical methods,
- B. those which are evaluated by other means.

There is not always a simple correspondence between the classification into categories A or B and the previously used classification into “random” and “systematic” uncertainties. The term “systematic uncertainty” can be misleading and should be avoided.

Any detailed report of the uncertainty should consist of a complete list of the components, specifying for each the method used to obtain its numerical value.

2. The components in category A are characterized by the estimated variances  $s_i^2$  (or the estimated “standard deviations”  $s_i$ ) and the number of degrees of freedom  $\nu_i$ . Where appropriate, the covariances should be given.

3. The components in category B should be characterized by quantities  $u_j^2$ , which may be considered as approximations to the corresponding variances, the existence of which is assumed. The quantities  $u_j^2$  may be treated like variances and the quantities  $u_j$  like standard deviations. Where appropriate, the covariances should be treated in a similar way.

4. The combined uncertainty should be characterized by the numerical value obtained by applying the usual method for the combination of variances. The combined uncertainty and its components should be expressed in the form of “standard deviations.”

5. If, for particular applications, it is necessary to multiply the combined uncertainty by a factor to obtain an overall uncertainty, the multiplying factor used must always be stated.

# GUIDE TO THE EXPRESSION OF UNCERTAINTY IN MEASUREMENT

## 1 Scope

**1.1** This *Guide* establishes general rules for evaluating and expressing uncertainty in measurement that can be followed at various levels of accuracy and in many fields – from the shop floor to fundamental research. Therefore, the principles of this *Guide* are intended to be applicable to a broad spectrum of measurements, including those required for:

- maintaining quality control and quality assurance in production;
- complying with and enforcing laws and regulations;
- conducting basic research, and applied research and development, in science and engineering;
- calibrating standards and instruments and performing tests throughout a national measurement system in order to achieve traceability to national standards;
- developing, maintaining, and comparing international and national physical reference standards, including reference materials.

**1.2** This *Guide* is primarily concerned with the expression of uncertainty in the measurement of a well-defined physical quantity – the measurand – that can be characterized by an essentially unique value. If the phenomenon of interest can be represented only as a distribution of values or is dependent on one or more parameters, such as time, then the measurands required for its description are the set of quantities describing that distribution or that dependence.

**1.3** This *Guide* is also applicable to evaluating and expressing the uncertainty associated with the conceptual design and theoretical analysis of experiments, methods of measurement, and complex components and systems. Because a measurement result and its uncertainty may be conceptual and based entirely on hypothetical data, the term “result of a measurement” as used in this *Guide* should be interpreted in this broader context.

**1.4** This *Guide* provides general rules for evaluating and expressing uncertainty in measurement rather than detailed, technology-specific instructions. Further, it does not discuss how the uncertainty of a particular measurement result, once evaluated, may be used for different purposes, for example, to draw conclusions about the compatibility of that result with other similar results, to establish tolerance limits in a manufacturing process, or to decide if a certain course of action may be safely undertaken. It may therefore be necessary to develop particular standards based on this *Guide* that deal with the problems peculiar to specific fields of measurement or with the various uses of quantitative expressions of uncertainty. These standards may be simplified versions of this *Guide* but should include the detail that is appropriate to the level of accuracy and complexity of the measurements and uses addressed.

NOTE – There may be situations in which the concept of uncertainty of measurement is believed not to be fully applicable, such as when the precision of a test method is determined (see reference [5], for example).

## 2 Definitions

### 2.1 General metrological terms

The definition of a number of general metrological terms relevant to this *Guide*, such as “measurable quantity,” “measurand,” and “error of measurement,” are given in annex B. These definitions are taken from the *International vocabulary of basic and general terms in metrology* (abbreviated VIM) [6]. In addition, annex C gives the definitions of a number of basic statistical terms taken mainly from International Standard ISO 3534-1 [7]. When one of these metrological or statistical terms (or a closely related term) is first used in the text, starting with clause 3, it is printed in boldface and the number of the subclause in which it is defined is given in parentheses.

Because of its importance to this *Guide*, the definition of the general metrological term “uncertainty of measurement” is given both in annex B and 2.2.3. The definitions of the most important terms specific to this *Guide* are given in 2.3.1 to 2.3.6. In all of these subclauses and in annexes B and C, the use of parentheses around certain words of some terms means that these words may be omitted if this is unlikely to cause confusion.

### 2.2 The term “uncertainty”

The concept of uncertainty is discussed further in clause 3 and annex D.

**2.2.1** The word “uncertainty” means doubt, and thus in its broadest sense “uncertainty of measurement” means doubt about the validity of the result of a measurement. 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.

**2.2.2** In this *Guide*, the word “uncertainty” without

adjectives refers both to the general concept of uncertainty and to any or all quantitative measures of that concept. When a specific measure is intended, appropriate adjectives are used.

**2.2.3** The formal definition of the term “uncertainty of measurement” developed for use in this *Guide* and in the current VIM [6] (VIM entry 3.9) is as follows:

#### **uncertainty (of measurement)**

parameter, associated with the result of a measurement, that characterizes the dispersion of the values that could reasonably be attributed to the measurand

#### NOTES

- 1 The parameter may be, for example, a standard deviation (or a given multiple of it), or the half-width of an interval having a stated level of confidence.
- 2 Uncertainty of measurement comprises, in general, many components. Some of these components may be evaluated from the statistical distribution of the results of series of measurements and can be characterized by experimental standard deviations. The other components, which also can be characterized by standard deviations, are evaluated from assumed probability distributions based on experience or other information.
- 3 It is understood that the result of the measurement is the best estimate of the value of the measurand, and that all components of uncertainty, including those arising from systematic effects, such as components associated with corrections and reference standards, contribute to the dispersion.

**2.2.4** The definition of uncertainty of measurement given in 2.2.3 is an operational one that focuses on the measurement result and its evaluated uncertainty. However, it is not inconsistent with other concepts of uncertainty of measurement, such as

- a measure of the possible error in the estimated value of the measurand as provided by the result of a measurement;

- an estimate characterizing the range of values within which the true value of a measurand lies (VIM, first edition, 1984, entry 3.09).

Although these two traditional concepts are valid as ideals, they focus on *unknowable* quantities: the “error” of the result of a measurement and the “true value” of the measurand (in contrast to its estimated value), respectively. Nevertheless, whichever *concept* of uncertainty is adopted, an uncertainty component is always *evaluated* using the same data and related information. (See also E.5.)

## 2.3 Terms specific to this Guide

In general, terms that are specific to this *Guide* are defined in the text when first introduced. However, the definitions of the most important of these terms are given here for easy reference.

NOTE – Further discussion related to these terms may be found as follows: for 2.3.2, see 3.3.3 and 4.2; for 2.3.3, see 3.3.3 and 4.3; for 2.3.4, see clause 5 and equations (10) and (13); and for 2.3.5 and 2.3.6, see clause 6.

### 2.3.1 standard uncertainty

uncertainty of the result of a measurement expressed as a standard deviation

### 2.3.2 Type A evaluation (of uncertainty)

method of evaluation of uncertainty by the statistical analysis of series of observations

### 2.3.3 Type B evaluation (of uncertainty)

method of evaluation of uncertainty by means other than the statistical analysis of series of observations

### 2.3.4 combined standard uncertainty

standard uncertainty of the result of a measurement when that result is obtained from the values of a number of other quantities, equal to the positive square root of a sum of terms, the terms being the variances or covariances of these other quantities weighted according to how the measurement result varies with changes in these quantities

### 2.3.5 expanded uncertainty

quantity defining an interval about the result of a measurement that may be expected to encompass a large fraction of the distribution of values that could reasonably be attributed to the measurand

#### NOTES

- 1 The fraction may be viewed as the coverage probability or level of confidence of the interval.
- 2 To associate a specific level of confidence with the interval defined by the expanded uncertainty requires explicit or implicit assumptions regarding the probability distribution characterized by the measurement result and its combined standard uncertainty. The level of confidence that may be attributed to this interval can be known only to the extent to which such assumptions may be justified.
- 3 Expanded uncertainty is termed *overall uncertainty* in paragraph 5 of Recommendation INC-1 (1980).

### 2.3.6 coverage factor

numerical factor used as a multiplier of the combined standard uncertainty in order to obtain an expanded uncertainty

NOTE – A coverage factor,  $k$ , is typically in the range 2 to 3.

## 3 Basic concepts

Additional discussion of basic concepts may be found in annex D, which focuses on the ideas of “true” value, error, and uncertainty and includes graphical illustrations of these concepts; and in annex E, which explores the motivation and statistical basis for Recommendation INC-1 (1980) upon which this *Guide* rests. Annex J is a glossary of the principal mathematical symbols used throughout the *Guide*.

### 3.1 Measurement

**3.1.1** The objective of a **measurement** (B.2.5) is to determine the **value** (B.2.2) of the **measurand** (B.2.9), that is, the value of the **particular quantity** (B.2.1, note 1) to be measured. A measurement therefore begins with an appropriate specification of the measurand, the **method of measurement** (B.2.7), and the **measurement procedure** (B.2.8).

NOTE – The term “true value” (see annex D) is not used in this *Guide* for the reasons given in D.3.5; the terms “value of a measurand” (or of a quantity) and “true value of a measurand” (or of a quantity) are viewed as equivalent.

**3.1.2** In general, the **result of a measurement** (B.2.11) is only an approximation or **estimate** (C.2.26) of the value of the measurand and thus is complete only when accompanied by a statement of the **uncertainty** (B.2.18) of that estimate.

**3.1.3** In practice, the required specification or definition of the measurand is dictated by the required **accuracy of measurement** (B.2.14). The measurand should be defined with sufficient completeness with respect to the required accuracy so that for all practical purposes associated with the measurement its value is unique. It is in this sense that the expression “value of the measurand” is used in this *Guide*.

EXAMPLE – If the length of a nominally one-metre long steel

bar is to be determined to micrometre accuracy, its specification should include the temperature and pressure at which the length is defined. Thus the measurand should be specified as, for example, the length of the bar at 25,00 °C and 101 325 Pa (plus any other defining parameters deemed necessary, such as the way the bar is to be supported). However, if the length is to be determined to only millimetre accuracy, its specification would not require a defining temperature or pressure or a value for any other defining parameter.

NOTE – Incomplete definition of the measurand can give rise to a component of uncertainty sufficiently large that it must be included in the evaluation of the uncertainty of the measurement result (see D.1.1, D.3.4, and D.6.2).

**3.1.4** In many cases, the result of a measurement is determined on the basis of series of observations obtained under **repeatability conditions** (B.2.15, note 1).

**3.1.5** Variations in repeated observations are assumed to arise because **influence quantities** (B.2.10) that can affect the measurement result are not held completely constant.

**3.1.6** The mathematical model of the measurement that transforms the set of repeated observations into the measurement result is of critical importance because, in addition to the observations, it generally includes various influence quantities that are inexactly known. This lack of knowledge contributes to the uncertainty of the measurement result, as do the variations of the repeated observations and any uncertainty associated with the mathematical model itself.

**3.1.7** This *Guide* treats the measurand as a scalar (a single quantity). Extension to a set of related measurands determined simultaneously in the same measurement requires replacing the scalar measurand and its **variance** (C.2.11, C.2.20, C.3.2) by a vector measurand and **covariance matrix** (C.3.5). Such a replacement is considered in this *Guide* only in the examples (see H.2, H.3, and H.4).

### 3.2 Errors, effects, and corrections

**3.2.1** In general, a measurement has imperfections that give rise to an **error** (B.2.19) in the measurement result. Traditionally, an error is viewed as having two components, namely, a **random** (B.2.21) component and a **systematic** (B.2.22) component.

NOTE – Error is an idealized concept and errors cannot be known exactly.

**3.2.2** Random error presumably arises from unpredictable or stochastic temporal and spatial variations of influence quantities. The effects of such variations, hereafter termed *random effects*, give rise to variations in repeated observations of the measurand. Although it is not possible to compensate for the random error of a measurement result, it can usually be reduced by increasing the number of observations; its **expectation or expected value** (C.2.9, C.3.1) is zero.

#### NOTES

- 1 The experimental standard deviation of the arithmetic mean or average of a series of observations (see 4.2.3) is *not* the random error of the mean, although it is so designated in some publications. It is instead a measure of the *uncertainty* of the mean due to random effects. The exact value of the error in the mean arising from these effects cannot be known.
- 2 In this *Guide*, great care is taken to distinguish between the terms “error” and “uncertainty.” They are not synonyms, but represent completely different concepts; they should not be confused with one another or misused.

**3.2.3** Systematic error, like random error, cannot be eliminated but it too can often be reduced. If a systematic error arises from a recognized effect of an influence quantity on a measurement result, hereafter termed a *systematic effect*, the effect can be quantified and, if it is significant in size relative to the required accuracy of the measurement, a **correction** (B.2.23) or **correction factor** (B.2.24) can be applied to compensate for the effect. It is assumed that, after correction, the expectation or expected value of the error arising from a systematic effect is zero.

NOTE – The uncertainty of a correction applied to a measurement result to compensate for a systematic effect is *not* the systematic error, often termed bias, in the measurement result due to the effect as it is sometimes called. It is instead a measure of the *uncertainty* of the result due to incomplete knowledge of the required value of the correction. The error arising from imperfect compensation of a systematic effect cannot be exactly known. The terms “error” and “uncertainty” should be used properly and care taken to distinguish between them.

**3.2.4** It is assumed that the result of a measurement has been corrected for all recognized significant systematic effects and that every effort has been made to identify such effects.

EXAMPLE – A correction due to the finite impedance of a voltmeter used to determine the potential difference (the measurand) across a high-impedance resistor is applied to reduce the systematic effect on the result of the measurement arising from the loading effect of the voltmeter. However, the values of the impedances of the voltmeter and resistor, which are used to estimate the value of the correction and which are obtained from other measurements, are themselves uncertain. These uncertainties are used to evaluate the component of the uncertainty of the potential difference determination arising from the correction and thus from the systematic effect due to the finite impedance of the voltmeter.

#### NOTES

- 1 Often, measuring instruments and systems are adjusted or calibrated using measurement standards and reference materials to eliminate systematic effects; however, the uncertainties associated with these standards and materials must still be taken into account.
- 2 The case where a correction for a known significant systematic effect is not applied is discussed in the note to 6.3.1 and in F.2.4.5.

### 3.3 Uncertainty

**3.3.1** The uncertainty of the result of a measurement reflects the lack of exact knowledge of the value of the measurand (see 2.2). The result of a measurement after correction for recognized systematic effects is still only an *estimate* of the value of the measurand because of the uncertainty arising from random effects and from imperfect correction of the result for systematic effects.

NOTE – The result of a measurement (after correction) can unknowably be very close to the value of the measurand (and hence have a negligible error) even though it may have a large uncertainty. Thus the uncertainty of the result of a measurement should not be confused with the remaining unknown error.

**3.3.2** In practice, there are many possible sources of uncertainty in a measurement, including:

- a) incomplete definition of the measurand;
- b) imperfect realization of the definition of the measurand;
- c) nonrepresentative sampling – the sample measured may not represent the defined measurand;
- d) inadequate knowledge of the effects of environmental conditions on the measurement or

imperfect measurement of environmental conditions;

- e) personal bias in reading analogue instruments;
- f) finite instrument resolution or discrimination threshold;
- g) inexact values of measurement standards and reference materials;
- h) inexact values of constants and other parameters obtained from external sources and used in the data-reduction algorithm;
- i) approximations and assumptions incorporated in the measurement method and procedure;
- j) variations in repeated observations of the measurand under apparently identical conditions.

These sources are not necessarily independent, and some of sources a) to i) may contribute to source j). Of course, an unrecognized systematic effect cannot be taken into account in the evaluation of the uncertainty of the result of a measurement but contributes to its error.

**3.3.3** Recommendation INC-1 (1980) of the Working Group on the Statement of Uncertainties groups uncertainty components into two categories based on their method of evaluation, “A” and “B” (see 0.7, 2.3.2, and 2.3.3). These categories apply to *uncertainty* and are not substitutes for the words “random” and “systematic.” The uncertainty of a correction for a known systematic effect may in some cases be obtained by a Type A evaluation while in other cases by a Type B evaluation, as may the uncertainty characterizing a random effect.

NOTE – In some publications uncertainty components are categorized as “random” and “systematic” and are associated with errors arising from random effects and known systematic effects, respectively. Such categorization of components of uncertainty can be ambiguous when generally applied. For example, a “random” component of uncertainty in one measurement may become a “systematic” component of uncertainty in another measurement in which the result of the first measurement is used as an input datum. Categorizing the *methods* of evaluating uncertainty components rather than the *components* themselves avoids such ambiguity. At the same time, it does not preclude collecting individual components that have been evaluated by the two different methods into designated groups to be used for a particular purpose (see 3.4.3).

**3.3.4** The purpose of the Type A and Type B classification is to indicate the two different ways of evaluating uncertainty components and is for convenience of discussion only; the classification is not meant to indicate that there is any difference in the nature of the

components resulting from the two types of evaluation. Both types of evaluation are based on **probability distributions** (C.2.3), and the uncertainty components resulting from either type are quantified by variances or standard deviations.

**3.3.5** The estimated variance  $u^2$  characterizing an uncertainty component obtained from a Type A evaluation is calculated from series of repeated observations and is the familiar statistically estimated variance  $s^2$  (see 4.2). The estimated **standard deviation** (C.2.12, C.2.21, C.3.3)  $u$ , the positive square root of  $u^2$ , is thus  $u = s$  and for convenience is sometimes called a *Type A standard uncertainty*. For an uncertainty component obtained from a Type B evaluation, the estimated variance  $u^2$  is evaluated using available knowledge (see 4.3), and the estimated standard deviation  $u$  is sometimes called a *Type B standard uncertainty*.

Thus a Type A standard uncertainty is obtained from a **probability density function** (C.2.5) derived from an observed **frequency distribution** (C.2.18), while a Type B standard uncertainty is obtained from an assumed probability density function based on the degree of belief that an event will occur [often called subjective **probability** (C.2.1)]. Both approaches employ recognized interpretations of probability.

NOTE – A Type B evaluation of an uncertainty component is usually based on a pool of comparatively reliable information (see 4.3.1).

**3.3.6** The standard uncertainty of the result of a measurement, when that result is obtained from the values of a number of other quantities, is termed *combined standard uncertainty* and denoted by  $u_c$ . It is the estimated standard deviation associated with the result and is equal to the positive square root of the combined variance obtained from all variance and **covariance** (C.3.4) components, however evaluated, using what is termed in this *Guide* the *law of propagation of uncertainty* (see clause 5).

**3.3.7** To meet the needs of some industrial and commercial applications, as well as requirements in the areas of health and safety, an *expanded uncertainty*  $U$  is obtained by multiplying the combined standard uncertainty  $u_c$  by a *coverage factor*  $k$ . The intended purpose of  $U$  is to provide an interval about the result of a measurement that may be expected to encompass a large fraction of the distribution of values that could reasonably be attributed

to the measurand. The choice of the factor  $k$ , which is usually in the range 2 to 3, is based on the coverage probability or level of confidence required of the interval (see clause 6).

NOTE – The coverage factor  $k$  is always to be stated, so that the standard uncertainty of the measured quantity can be recovered for use in calculating the combined standard uncertainty of other measurement results that may depend on that quantity.

### 3.4 Practical considerations

**3.4.1** If all of the quantities on which the result of a measurement depends are varied, its uncertainty can be evaluated by statistical means. However, because this is rarely possible in practice due to limited time and resources, the uncertainty of a measurement result is usually evaluated using a mathematical model of the measurement and the law of propagation of uncertainty. Thus implicit in this *Guide* is the assumption that a measurement can be modelled mathematically to the degree imposed by the required accuracy of the measurement.

**3.4.2** Because the mathematical model may be incomplete, all relevant quantities should be varied to the fullest practicable extent so that the evaluation of uncertainty can be based as much as possible on observed data. Whenever feasible, the use of empirical models of the measurement founded on long-term quantitative data, and the use of check standards and control charts that can indicate if a measurement is under statistical control, should be part of the effort to obtain reliable evaluations of uncertainty. The mathematical model should always be revised when the observed data, including the result of independent determinations of the same measurand, demonstrate that the model is incomplete. A well-designed experiment can greatly facilitate reliable evaluations of uncertainty and is an important part of the art of measurement.

**3.4.3** In order to decide if a measurement system is functioning properly, the experimentally observed variability of its output values, as measured by their observed standard deviation, is often compared with the predicted standard deviation obtained by combining the various uncertainty components that characterize the measurement. In such cases, only those components (whether obtained from Type A or Type B evaluations) that could contribute to the experimentally observed variability of these output values should be considered.

NOTE – Such an analysis may be facilitated by gathering those components that contribute to the variability and those that do not into two separate and appropriately labelled groups.

**3.4.4** In some cases, the uncertainty of a correction for a systematic effect need not be included in the evaluation of the uncertainty of a measurement result. Although the uncertainty has been evaluated, it may be ignored if its contribution to the combined standard uncertainty of the measurement result is insignificant. If the value of the correction itself is insignificant relative to the combined standard uncertainty, it too may be ignored.

**3.4.5** It often occurs in practice, especially in the domain of legal metrology, that a device is tested through a comparison with a measurement standard and the uncertainties associated with the standard and the comparison procedure are negligible relative to the required accuracy of the test. An example is the use of a set of well-calibrated standards of mass to test the accuracy of a commercial scale. In such cases, because the components of uncertainty are small enough to be ignored, the measurement may be viewed as determining the error of the device under test. (See also F.2.4.2.)

**3.4.6** The estimate of the value of a measurand provided by the result of a measurement is sometimes expressed in terms of the adopted value of a measurement standard rather than in terms of the relevant unit of the International System of Units (SI). In such cases the magnitude of the uncertainty ascribable to the measurement result may be significantly smaller than when that result is expressed in the relevant SI unit. (In effect, the measurand has been redefined to be the ratio of the value of the quantity to be measured to the adopted value of the standard.)

EXAMPLE – A high-quality Zener voltage standard is calibrated by comparison with a Josephson effect voltage reference based on the conventional value of the Josephson constant recommended for international use by the CIPM. The relative combined standard uncertainty  $u_c(V_S)/V_S$  (see 5.1.6) of the calibrated potential difference  $V_S$  of the Zener standard is  $2 \times 10^{-8}$  when  $V_S$  is reported in terms of the conventional value, but  $u_c(V_S)/V_S$  is  $4 \times 10^{-7}$  when  $V_S$  is reported in terms of the SI unit of potential difference, the volt (V), because of the additional uncertainty associated with the SI value of the Josephson constant.

**3.4.7** Blunders in recording or analysing data can introduce a significant unknown error in the result of a measurement. Large blunders can usually be identified by

a proper review of the data; small ones could be masked by, or even appear as, random variations. Measures of uncertainty are not intended to account for such mistakes.

**3.4.8** Although this *Guide* provides a framework for assessing uncertainty, it cannot substitute for critical thinking, intellectual honesty, and professional skill. The

evaluation of uncertainty is neither a routine task nor a purely mathematical one; it depends on detailed knowledge of the nature of the measurand and of the measurement. The quality and utility of the uncertainty quoted for the result of a measurement therefore ultimately depend on the understanding, critical analysis, and integrity of those who contribute to the assignment of its value.

## 4 Evaluating standard uncertainty

Additional guidance on evaluating uncertainty components, mainly of a practical nature, may be found in annex F.

### 4.1 Modelling the measurement

**4.1.1** In most cases a measurand  $Y$  is not measured directly, but is determined from  $N$  other quantities  $X_1, X_2, \dots, X_N$  through a functional relationship  $f$ :

$$Y = f(X_1, X_2, \dots, X_N) \quad \dots (1)$$

#### NOTES

1 For economy of notation, in this *Guide* the same symbol is used for the physical quantity (the measurand) and for the random variable (see 4.2.1) that represents the possible outcome of an observation of that quantity. When it is stated that  $X_i$  has a particular probability distribution, the symbol is used in the latter sense; it is assumed that the physical quantity itself can be characterized by an essentially unique value (see 1.2 and 3.1.3).

2 In a series of observations, the  $k$ th observed value of  $X_i$  is denoted by  $X_{i,k}$ ; hence if  $R$  denotes the resistance of a resistor, the  $k$ th observed value of the resistance is denoted by  $R_k$ .

3 The estimate of  $X_i$  (strictly speaking, of its expectation) is denoted by  $x_i$ .

**EXAMPLE** – If a potential difference  $V$  is applied to the terminals of a temperature-dependent resistor that has a resistance  $R_0$  at the defined temperature  $t_0$  and a linear temperature coefficient of resistance  $\alpha$ , the power  $P$  (the measurand) dissipated by the resistor at the temperature  $t$  depends on  $V, R_0, \alpha$ , and  $t$  according to

$$P = f(V, R_0, \alpha, t) = V^2 / R_0 [1 + \alpha(t - t_0)]$$

**NOTE** – Other methods of measuring  $P$  would be modelled by different mathematical expressions.

**4.1.2** The input quantities  $X_1, X_2, \dots, X_N$  upon which the output quantity  $Y$  depends may themselves be viewed as measurands and may themselves depend on other

quantities, including corrections and correction factors for systematic effects, thereby leading to a complicated functional relationship  $f$  that may never be written down explicitly. Further,  $f$  may be determined experimentally (see 5.1.4) or exist only as an algorithm that must be evaluated numerically. The function  $f$  as it appears in this *Guide* is to be interpreted in this broader context, in particular as that function which contains every quantity, including all corrections and correction factors, that can contribute a significant component of uncertainty to the measurement result.

Thus, if data indicate that  $f$  does not model the measurement to the degree imposed by the required accuracy of the measurement result, additional input quantities must be included in  $f$  to eliminate the inadequacy (see 3.4.2). This may require introducing an input quantity to reflect incomplete knowledge of a phenomenon that affects the measurand. In the example of 4.1.1, additional input quantities might be needed to account for a known nonuniform temperature distribution across the resistor, a possible nonlinear temperature coefficient of resistance, or a possible dependence of resistance on barometric pressure.

**NOTE** – Nonetheless, equation (1) may be as elementary as  $Y = X_1 - X_2$ . This expression models, for example, the comparison of two determinations of the same quantity  $X$ .

**4.1.3** The set of input quantities  $X_1, X_2, \dots, X_N$  may be categorized as:

- quantities whose values *and uncertainties* are directly determined in the current measurement. These values and uncertainties may be obtained from, for example, a single observation, repeated observations, or judgement based on experience, and may involve the determination of corrections to instrument readings and corrections for influence

quantities, such as ambient temperature, barometric pressure, and humidity;

- quantities whose values and uncertainties are brought into the measurement from external sources, such as quantities associated with calibrated measurement standards, certified reference materials, and reference data obtained from handbooks.

**4.1.4** An estimate of the measurand  $Y$ , denoted by  $y$ , is obtained from equation (1) using *input estimates*  $x_1, x_2, \dots, x_N$  for the values of the  $N$  quantities  $X_1, X_2, \dots, X_N$ . Thus the *output estimate*  $y$ , which is the result of the measurement, is given by

$$y = f(x_1, x_2, \dots, x_N) \quad \dots (2)$$

NOTE – In some cases the estimate  $y$  may be obtained from

$$y = \bar{Y} = \frac{1}{n} \sum_{k=1}^n Y_k = \frac{1}{n} \sum_{k=1}^n f(X_{1,k}, X_{2,k}, \dots, X_{N,k})$$

That is,  $y$  is taken as the arithmetic mean or average (see 4.2.1) of  $n$  independent determinations  $Y_k$  of  $Y$ , each determination having the same uncertainty and each being based on a complete set of observed values of the  $N$  input quantities  $X_i$  obtained at the same time. This way of averaging, rather than  $y = f(\bar{X}_1, \bar{X}_2, \dots, \bar{X}_N)$ , where  $\bar{X}_i = (\sum_{k=1}^n X_{i,k})/n$  is the arithmetic mean of the individual observations  $X_{i,k}$ , may be preferable when  $f$  is a nonlinear function of the input quantities  $X_1, X_2, \dots, X_N$ , but the two approaches are identical if  $f$  is a linear function of the  $X_i$  (see H.2 and H.4).

**4.1.5** The estimated standard deviation associated with the output estimate or measurement result  $y$ , termed *combined standard uncertainty* and denoted by  $u_c(y)$ , is determined from the estimated standard deviation associated with each input estimate  $x_i$ , termed *standard uncertainty* and denoted by  $u(x_i)$  (see 3.3.5 and 3.3.6).

**4.1.6** Each input estimate  $x_i$  and its associated standard uncertainty  $u(x_i)$  are obtained from a distribution of possible values of the input quantity  $X_i$ . This probability distribution may be frequency based, that is, based on a series of observations  $X_{i,k}$  of  $X_i$ , or it may be an *a priori* distribution. Type A evaluations of standard uncertainty components are founded on frequency distributions while Type B evaluations are founded on *a priori* distributions. It must be recognized that in both cases the distributions are models that are used to represent the state of our knowledge.

## 4.2 Type A evaluation of standard uncertainty

**4.2.1** In most cases, the best available estimate of the expectation or expected value  $\mu_q$  of a quantity  $q$  that varies randomly [a **random variable** (C.2.2)], and for which  $n$  independent observations  $q_k$  have been obtained under the same conditions of measurement (see B.2.15), is the **arithmetic mean** or **average**  $\bar{q}$  (C.2.19) of the  $n$  observations:

$$\bar{q} = \frac{1}{n} \sum_{k=1}^n q_k \quad \dots (3)$$

Thus, for an input quantity  $X_i$  estimated from  $n$  independent repeated observations  $X_{i,k}$ , the arithmetic mean  $\bar{X}_i$  obtained from equation (3) is used as the input estimate  $x_i$  in equation (2) to determine the measurement result  $y$ ; that is,  $x_i = \bar{X}_i$ . Those input estimates not evaluated from repeated observations must be obtained by other methods, such as those indicated in the second category of 4.1.3.

**4.2.2** The individual observations  $q_k$  differ in value because of random variations in the influence quantities, or random effects (see 3.2.2). The experimental variance of the observations, which estimates the variance  $\sigma^2$  of the probability distribution of  $q$ , is given by

$$s^2(q_k) = \frac{1}{n-1} \sum_{k=1}^n (q_k - \bar{q})^2 \quad \dots (4)$$

This estimate of variance and its positive square root  $s(q_k)$ , termed the **experimental standard deviation** (B.2.17), characterize the variability of the observed values  $q_k$ , or more specifically, their dispersion about their mean  $\bar{q}$ .

**4.2.3** The best estimate of  $\sigma^2(\bar{q}) = \sigma^2/n$ , the variance of the mean, is given by

$$s^2(\bar{q}) = \frac{s^2(q_k)}{n} \quad \dots (5)$$

The experimental variance of the mean  $s^2(\bar{q})$  and the **experimental standard deviation of the mean**  $s(\bar{q})$  (B.2.17, note 2), equal to the positive square root of  $s^2(\bar{q})$ , quantify how well  $\bar{q}$  estimates the expectation  $\mu_q$  of  $q$ , and either may be used as a measure of the uncertainty of  $\bar{q}$ .

Thus, for an input quantity  $X_i$  determined from  $n$  independent repeated observations  $X_{i,k}$ , the standard

uncertainty  $u(x_i)$  of its estimate  $x_i = \bar{X}_i$  is  $u(x_i) = s(\bar{X}_i)$ , with  $s^2(\bar{X}_i)$  calculated according to equation (5). For convenience,  $u^2(x_i) = s^2(\bar{X}_i)$  and  $u(x_i) = s(\bar{X}_i)$  are sometimes called a *Type A variance* and a *Type A standard uncertainty*, respectively.

#### NOTES

1 The number of observations  $n$  should be large enough to ensure that  $\bar{q}$  provides a reliable estimate of the expectation  $\mu_q$  of the random variable  $q$  and that  $s^2(\bar{q})$  provides a reliable estimate of the variance  $\sigma^2(\bar{q}) = \sigma^2/n$  (see 4.3.2, note). The difference between  $s^2(\bar{q})$  and  $\sigma^2(\bar{q})$  must be considered when one constructs confidence intervals (see 6.2.2). In this case, if the probability distribution of  $q$  is a normal distribution (see 4.3.4), the difference is taken into account through the  $t$ -distribution (see G.3.2).

2 Although the variance  $s^2(\bar{q})$  is the more fundamental quantity, the standard deviation  $s(\bar{q})$  is more convenient in practice because it has the same dimension as  $q$  and a more easily comprehended value than that of the variance.

**4.2.4** For a well-characterized measurement under statistical control, a combined or pooled estimate of variance  $s_p^2$  (or a pooled experimental standard deviation  $s_p$ ) that characterizes the measurement may be available. In such cases, when the value of a measurand  $q$  is determined from  $n$  independent observations, the experimental variance of the arithmetic mean  $\bar{q}$  of the observations is estimated better by  $s_p^2/n$  than by  $s^2(\bar{q})/n$  and the standard uncertainty is  $u = s_p/\sqrt{n}$ . (See also the note to H.3.6.)

**4.2.5** Often an estimate  $x_i$  of an input quantity  $X_i$  is obtained from a curve that has been fitted to experimental data by the method of least squares. The estimated variances and resulting standard uncertainties of the fitted parameters characterizing the curve and of any predicted points can usually be calculated by well-known statistical procedures (see H.3 and reference [8]).

**4.2.6** The **degrees of freedom** (C.2.31)  $\nu_i$  of  $u(x_i)$  (see G.3), equal to  $n - 1$  in the simple case where  $x_i = \bar{X}_i$  and  $u(x_i) = s(\bar{X}_i)$  are calculated from  $n$  independent observations as in 4.2.1 and 4.2.3, should always be given when Type A evaluations of uncertainty components are documented.

**4.2.7** If the random variations in the observations of an input quantity are correlated, for example, in time, the mean and experimental standard deviation of the mean as given in 4.2.1 and 4.2.3 may be inappropriate **estimators**

(C.2.25) of the desired **statistics** (C.2.23). In such cases, the observations should be analysed by statistical methods specially designed to treat a series of correlated, randomly-varying measurements.

NOTE – Such specialized methods are used to treat measurements of frequency standards. However, it is possible that as one goes from short-term measurements to long-term measurements of other metrological quantities, the assumption of uncorrelated random variations may no longer be valid and the specialized methods could be used to treat these measurements as well. (See reference [9], for example, for a detailed discussion of the Allan variance.)

**4.2.8** The discussion of Type A evaluation of standard uncertainty in 4.2.1 to 4.2.7 is not meant to be exhaustive; there are many situations, some rather complex, that can be treated by statistical methods. An important example is the use of calibration designs, often based on the method of least squares, to evaluate the uncertainties arising from both short- and long-term random variations in the results of comparisons of material artefacts of unknown values, such as gauge blocks and standards of mass, with reference standards of known values. In such comparatively simple measurement situations, components of uncertainty can frequently be evaluated by the statistical analysis of data obtained from designs consisting of nested sequences of measurements of the measurand for a number of different values of the quantities upon which it depends – a so-called analysis of variance (see H.5).

NOTE – At lower levels of the calibration chain, where reference standards are often assumed to be exactly known because they have been calibrated by a national or primary standards laboratory, the uncertainty of a calibration result may be a single Type A standard uncertainty evaluated from the pooled experimental standard deviation that characterizes the measurement.

### 4.3 Type B evaluation of standard uncertainty

**4.3.1** For an estimate  $x_i$  of an input quantity  $X_i$  that has not been obtained from repeated observations, the associated estimated variance  $u^2(x_i)$  or the standard uncertainty  $u(x_i)$  is evaluated by scientific judgement based on all of the available information on the possible variability of  $X_i$ . The pool of information may include

- previous measurement data;
- experience with or general knowledge of the behaviour and properties of relevant materials and

- instruments;
- manufacturer's specifications;
  - data provided in calibration and other certificates;
  - uncertainties assigned to reference data taken from handbooks.

For convenience,  $u^2(x_i)$  and  $u(x_i)$  evaluated in this way are sometimes called a *Type B variance* and a *Type B standard uncertainty*, respectively.

NOTE – When  $x_i$  is obtained from an *a priori* distribution, the associated variance is appropriately written as  $u^2(X_i)$ , but for simplicity,  $u^2(x_i)$  and  $u(x_i)$  are used throughout this *Guide*.

**4.3.2** The proper use of the pool of available information for a Type B evaluation of standard uncertainty calls for insight based on experience and general knowledge, and is a skill that can be learned with practice. It should be recognized that a Type B evaluation of standard uncertainty can be as reliable as a Type A evaluation, especially in a measurement situation where a Type A evaluation is based on a comparatively small number of statistically independent observations.

NOTE – If the probability distribution of  $q$  in note 1 to 4.2.3 is normal, then  $\sigma[s(\bar{q})]/\sigma(\bar{q})$ , the standard deviation of  $s(\bar{q})$  relative to  $\sigma(\bar{q})$ , is approximately  $[2(n-1)]^{-1/2}$ . Thus, taking  $\sigma[s(\bar{q})]$  as the uncertainty of  $s(\bar{q})$ , for  $n = 10$  observations the relative uncertainty in  $s(\bar{q})$  is 24 percent, while for  $n = 50$  observations it is 10 percent. (Additional values are given in table E.1 in annex E.)

**4.3.3** If the estimate  $x_i$  is taken from a manufacturer's specification, calibration certificate, handbook, or other source and its quoted uncertainty is stated to be a particular multiple of a standard deviation, the standard uncertainty  $u(x_i)$  is simply the quoted value divided by the multiplier, and the estimated variance  $u^2(x_i)$  is the square of that quotient.

EXAMPLE – A calibration certificate states that the mass of a stainless steel mass standard  $m_S$  of nominal value one kilogram is 1 000,000 325 g and that “the uncertainty of this value is 240  $\mu$ g at the three standard deviation level.” The standard uncertainty of the mass standard is then simply  $u(m_S) = (240 \mu\text{g})/3 = 80 \mu\text{g}$ . This corresponds to a relative standard uncertainty  $u(m_S)/m_S$  of  $80 \times 10^{-9}$  (see 5.1.6). The estimated variance is  $u^2(m_S) = (80 \mu\text{g})^2 = 6,4 \times 10^{-9} \text{ g}^2$ .

NOTE – In many cases little or no information is provided about the individual components from which the quoted uncertainty has been obtained. This is generally unimportant for expressing uncertainty according to the practices of this *Guide* since all standard uncertainties are treated in the same way when the

combined standard uncertainty of a measurement result is calculated (see clause 5).

**4.3.4** The quoted uncertainty of  $x_i$  is not necessarily given as a multiple of a standard deviation as in 4.3.3. Instead, one may find it stated that the quoted uncertainty defines an interval having a 90, 95, or 99 percent level of confidence (see 6.2.2). Unless otherwise indicated, one may assume that a **normal distribution** (C.2.14) was used to calculate the quoted uncertainty, and recover the standard uncertainty of  $x_i$  by dividing the quoted uncertainty by the appropriate factor for the normal distribution. The factors corresponding to the above three levels of confidence are 1,64; 1,96; and 2,58 (see also table G.1 in annex G).

NOTE – There would be no need for such an assumption if the uncertainty had been given in accordance with the recommendations of this *Guide* regarding the reporting of uncertainty, which stress that the coverage factor used is always to be given (see 7.2.3).

EXAMPLE – A calibration certificate states that the resistance of a standard resistor  $R_S$  of nominal value ten ohms is  $10,000\,742 \Omega \pm 129 \mu\Omega$  at  $23^\circ\text{C}$  and that “the quoted uncertainty of  $129 \mu\Omega$  defines an interval having a level of confidence of 99 percent.” The standard uncertainty of the resistor may be taken as  $u(R_S) = (129 \mu\Omega)/2,58 = 50 \mu\Omega$ , which corresponds to a relative standard uncertainty  $u(R_S)/R_S$  of  $5,0 \times 10^{-6}$  (see 5.1.6). The estimated variance is  $u^2(R_S) = (50 \mu\Omega)^2 = 2,5 \times 10^{-9} \Omega^2$ .

**4.3.5** Consider the case where, based on the available information, one can state that “there is a fifty-fifty chance that the value of the input quantity  $X_i$  lies in the interval  $a_-$  to  $a_+$ ” (in other words, the probability that  $X_i$  lies within this interval is 0,5 or 50 percent). If it can be assumed that the distribution of possible values of  $X_i$  is approximately normal, then the best estimate  $x_i$  of  $X_i$  can be taken to be the midpoint of the interval. Further, if the half-width of the interval is denoted by  $a = (a_+ - a_-)/2$ , one can take  $u(x_i) = 1,48a$ , because for a normal distribution with expectation  $\mu$  and standard deviation  $\sigma$  the interval  $\mu \pm \sigma/1,48$  encompasses approximately 50 percent of the distribution.

EXAMPLE – A machinist determining the dimensions of a part estimates that its length lies, with probability 0,5, in the interval 10,07 mm to 10,15 mm, and reports that  $l = (10,11 \pm 0,04)$  mm, meaning that  $\pm 0,04$  mm defines an interval having a level of confidence of 50 percent. Then  $a = 0,04$  mm, and if one assumes a normal distribution for the possible values of  $l$ , the standard uncertainty of the length is  $u(l) = 1,48 \times 0,04 \text{ mm} \approx$

0,06 mm and the estimated variance is  $u^2(l) = (1,48 \times 0,04 \text{ mm})^2 = 3,5 \times 10^{-3} \text{ mm}^2$ .

**4.3.6** Consider a case similar to that of 4.3.5 but where, based on the available information, one can state that “there is about a two out of three chance that the value of  $X_i$  lies in the interval  $a_-$  to  $a_+$ ” (in other words, the probability that  $X_i$  lies within this interval is about 0,67). One can then reasonably take  $u(x_i) = a$ , because for a normal distribution with expectation  $\mu$  and standard deviation  $\sigma$  the interval  $\mu \pm \sigma$  encompasses about 68,3 percent of the distribution.

NOTE – It would give the value of  $u(x_i)$  considerably more significance than is obviously warranted if one were to use the actual normal deviate 0,96742 corresponding to probability  $p = 2/3$ , that is, if one were to write  $u(x_i) = a/0,96742 = 1,033a$ .

**4.3.7** In other cases it may be possible to estimate only bounds (upper and lower limits) for  $X_i$ , in particular, to state that “the probability that the value of  $X_i$  lies within the interval  $a_-$  to  $a_+$  for all practical purposes is equal to one and the probability that  $X_i$  lies outside this interval is essentially zero.” If there is *no specific knowledge* about the possible values of  $X_i$  within the interval, one can only assume that it is equally probable for  $X_i$  to lie anywhere within it (a uniform or rectangular distribution of possible values – see 4.4.5 and figure 2a). Then  $x_i$ , the expectation or expected value of  $X_i$ , is the midpoint of the interval,  $x_i = (a_- + a_+)/2$ , with associated variance

$$u^2(x_i) = (a_+ - a_-)^2 / 12 \quad \dots (6)$$

If the difference between the bounds,  $a_+ - a_-$ , is denoted by  $2a$ , then equation (6) becomes

$$u^2(x_i) = a^2 / 3 \quad \dots (7)$$

NOTE – When a component of uncertainty determined in this manner contributes significantly to the uncertainty of a measurement result, it is prudent to obtain additional data for its further evaluation.

#### EXAMPLES

1 A handbook gives the value of the coefficient of linear thermal expansion of pure copper at 20 °C,  $\alpha_{20}(\text{Cu})$ , as  $16,52 \times 10^{-6} \text{ }^\circ\text{C}^{-1}$  and simply states that “the error in this value should not exceed  $0,40 \times 10^{-6} \text{ }^\circ\text{C}^{-1}$ .” Based on this limited information, it is not unreasonable to assume that the value of  $\alpha_{20}(\text{Cu})$  lies with equal probability in the interval  $16,12 \times 10^{-6} \text{ }^\circ\text{C}^{-1}$  to  $16,92 \times 10^{-6} \text{ }^\circ\text{C}^{-1}$ , and that it is very unlikely that  $\alpha_{20}(\text{Cu})$  lies outside this interval. The variance of

this symmetric rectangular distribution of possible values of  $\alpha_{20}(\text{Cu})$  of half-width  $a = 0,40 \times 10^{-6} \text{ }^\circ\text{C}^{-1}$  is then, from equation (7),  $u^2(\alpha_{20}) = (0,40 \times 10^{-6} \text{ }^\circ\text{C}^{-1})^2 / 3 = 53,3 \times 10^{-15} \text{ }^\circ\text{C}^{-2}$ , and the standard uncertainty is  $u(\alpha_{20}) = (0,40 \times 10^{-6} \text{ }^\circ\text{C}^{-1}) / \sqrt{3} = 0,23 \times 10^{-6} \text{ }^\circ\text{C}^{-1}$ .

2 A manufacturer’s specifications for a digital voltmeter state that “between one and two years after the instrument is calibrated, its accuracy on the 1 V range is  $14 \times 10^{-6}$  times the reading plus  $2 \times 10^{-6}$  times the range.” Consider that the instrument is used 20 months after calibration to measure on its 1 V range a potential difference  $V$ , and the arithmetic mean of a number of independent repeated observations of  $V$  is found to be  $\bar{V} = 0,928\,571 \text{ V}$  with a Type A standard uncertainty  $u(\bar{V}) = 12 \text{ } \mu\text{V}$ . One can obtain the standard uncertainty associated with the manufacturer’s specifications from a Type B evaluation by assuming that the stated accuracy provides symmetric bounds to an additive correction to  $\bar{V}$ ,  $\Delta\bar{V}$ , of expectation equal to zero and with equal probability of lying anywhere within the bounds. The half-width  $a$  of the symmetric rectangular distribution of possible values of  $\Delta\bar{V}$  is then  $a = (14 \times 10^{-6}) \times (0,928\,571 \text{ V}) + (2 \times 10^{-6}) \times (1 \text{ V}) = 15 \text{ } \mu\text{V}$ , and from equation (7),  $u^2(\Delta\bar{V}) = 75 \text{ } \mu\text{V}^2$  and  $u(\Delta\bar{V}) = 8,7 \text{ } \mu\text{V}$ . The estimate of the value of the measurand  $V$ , for simplicity denoted by the same symbol  $V$ , is given by  $V = \bar{V} + \Delta\bar{V} = 0,928\,571 \text{ V}$ . One can obtain the combined standard uncertainty of this estimate by combining the 12  $\mu\text{V}$  Type A standard uncertainty of  $\bar{V}$  with the 8,7  $\mu\text{V}$  Type B standard uncertainty of  $\Delta\bar{V}$ . The general method for combining standard uncertainty components is given in clause 5, with this particular example treated in 5.1.5.

**4.3.8** In 4.3.7 the upper and lower bounds  $a_+$  and  $a_-$  for the input quantity  $X_i$  may not be symmetric with respect to its best estimate  $x_i$ ; more specifically, if the lower bound is written as  $a_- = x_i - b_-$  and the upper bound as  $a_+ = x_i + b_+$ , then  $b_- \neq b_+$ . Since in this case  $x_i$  (assumed to be the expectation of  $X_i$ ) is not at the centre of the interval  $a_-$  to  $a_+$ , the probability distribution of  $X_i$  cannot be uniform throughout the interval. However, there may not be enough information available to choose an appropriate distribution; different models will lead to different expressions for the variance. In the absence of such information the simplest approximation is

$$u^2(x_i) = \frac{(b_+ + b_-)^2}{12} = \frac{(a_+ - a_-)^2}{12} \quad \dots (8)$$

which is the variance of a rectangular distribution with full width  $b_+ + b_-$ . (Asymmetric distributions are also discussed in F.2.4.4 and G.5.3.)

EXAMPLE – If in example 1 of 4.3.7 the value of the coefficient is given in the handbook as  $\alpha_{20}(\text{Cu}) = 16,52 \times$

$10^{-6} \text{ }^{\circ}\text{C}^{-1}$  and it is stated that “the smallest possible value is  $16,40 \times 10^{-6} \text{ }^{\circ}\text{C}^{-1}$  and the largest possible value is  $16,92 \times 10^{-6} \text{ }^{\circ}\text{C}^{-1}$ ,” then  $b_- = 0,12 \times 10^{-6} \text{ }^{\circ}\text{C}^{-1}$ ,  $b_+ = 0,40 \times 10^{-6} \text{ }^{\circ}\text{C}^{-1}$ , and, from equation (8),  $u(a_{20}) = 0,15 \times 10^{-6} \text{ }^{\circ}\text{C}^{-1}$ .

## NOTES

1 In many practical measurement situations where the bounds are asymmetric, it may be appropriate to apply a correction to the estimate  $x_i$  of magnitude  $(b_+ - b_-)/2$  so that the new estimate  $x'_i$  of  $X_i$  is at the midpoint of the bounds:  $x'_i = (a_- + a_+)/2$ . This reduces the situation to the case of 4.3.7, with new values  $b'_+ = b'_- = (b_+ + b_-)/2 = (a_+ - a_-)/2 = a$ .

2 Based on the principle of maximum entropy, the probability density function in the asymmetric case may be shown to be  $p(X_i) = A \exp[-\lambda(X_i - x_i)]$ , with  $A = [b_- \exp(\lambda b_-) + b_+ \exp(-\lambda b_+)]^{-1}$  and  $\lambda = \{\exp[\lambda(b_- + b_+)] - 1\} / \{b_- \exp[\lambda(b_- + b_+)] + b_+\}$ . This leads to the variance  $u^2(x_i) = b_+ b_- - (b_+ - b_-)^2 / \lambda$ ; for  $b_+ > b_-$ ,  $\lambda > 0$  and for  $b_+ < b_-$ ,  $\lambda < 0$ .

**4.3.9** In 4.3.7, because there was no specific knowledge about the possible values of  $X_i$  within its estimated bounds  $a_-$  to  $a_+$ , one could only assume that it was equally probable for  $X_i$  to take any value within those bounds, with zero probability of being outside them. Such step function discontinuities in a probability distribution are often unphysical. In many cases it is more realistic to expect that values near the bounds are less likely than those near the midpoint. It is then reasonable to replace the symmetric rectangular distribution with a symmetric trapezoidal distribution having equal sloping sides (an isosceles trapezoid), a base of width  $a_+ - a_- = 2a$ , and a top of width  $2a\beta$ , where  $0 \leq \beta \leq 1$ . As  $\beta \rightarrow 1$  this trapezoidal distribution approaches the rectangular distribution of 4.3.7, while for  $\beta = 0$  it is a triangular distribution (see 4.4.6 and figure 2b). Assuming such a trapezoidal distribution for  $X_i$ , one finds that the expectation of  $X_i$  is  $x_i = (a_- + a_+)/2$  and its associated variance is

$$u^2(x_i) = a^2(1 + \beta^2)/6 \quad \dots (9a)$$

which becomes for the triangular distribution,  $\beta = 0$ ,

$$u^2(x_i) = a^2/6 \quad \dots (9b)$$

## NOTES

1 For a normal distribution with expectation  $\mu$  and standard deviation  $\sigma$ , the interval  $\mu \pm 3\sigma$  encompasses approximately 99,73 percent of the distribution. Thus, if the upper and lower bounds  $a_+$  and  $a_-$  define 99,73 percent limits rather than 100

percent limits, and  $X_i$  can be assumed to be approximately normally distributed rather than there being no specific knowledge about  $X_i$  between the bounds as in 4.3.7, then  $u^2(x_i) = a^2/9$ . By comparison, the variance of a symmetric rectangular distribution of half-width  $a$  is  $a^2/3$  [equation (7)] and that of a symmetric triangular distribution of half width  $a$  is  $a^2/6$  [equation (9b)]. The magnitudes of the variances of the three distributions are surprisingly similar in view of the large differences in the amount of information required to justify them.

2 The trapezoidal distribution is equivalent to the convolution of two rectangular distributions [10], one with a half-width  $a_1$  equal to the mean half-width of the trapezoid,  $a_1 = a(1 + \beta)/2$ , the other with a half-width  $a_2$  equal to the mean width of one of the triangular portions of the trapezoid,  $a_2 = a(1 - \beta)/2$ . The variance of the distribution is  $u^2 = a_1^2/3 + a_2^2/3$ . The convolved distribution can be interpreted as a rectangular distribution whose width  $2a_1$  has itself an uncertainty represented by a rectangular distribution of width  $2a_2$  and models the fact that the bounds on an input quantity are not exactly known. But even if  $a_2$  is as large as 30 percent of  $a_1$ ,  $u$  exceeds  $a_1/\sqrt{3}$  by less than 5 percent.

**4.3.10** It is important not to “double-count” uncertainty components. If a component of uncertainty arising from a particular effect is obtained from a Type B evaluation, it should be included as an independent component of uncertainty in the calculation of the combined standard uncertainty of the measurement result only to the extent that the effect does not contribute to the observed variability of the observations. This is because the uncertainty due to that portion of the effect that contributes to the observed variability is already included in the component of uncertainty obtained from the statistical analysis of the observations.

**4.3.11** The discussion of Type B evaluation of standard uncertainty in 4.3.3 to 4.3.9 is meant only to be indicative. Further, evaluations of uncertainty should be based on quantitative data to the maximum extent possible, as emphasized in 3.4.1 and 3.4.2.

## 4.4 Graphical illustration of evaluating standard uncertainty

**4.4.1** Figure 1 represents the estimation of the value of an input quantity  $X_i$  and the evaluation of the uncertainty of that estimate from the unknown distribution of possible measured values of  $X_i$ , or probability distribution of  $X_i$ , that is sampled by means of repeated observations.

**4.4.2** In figure 1a it is assumed that the input quantity  $X_i$

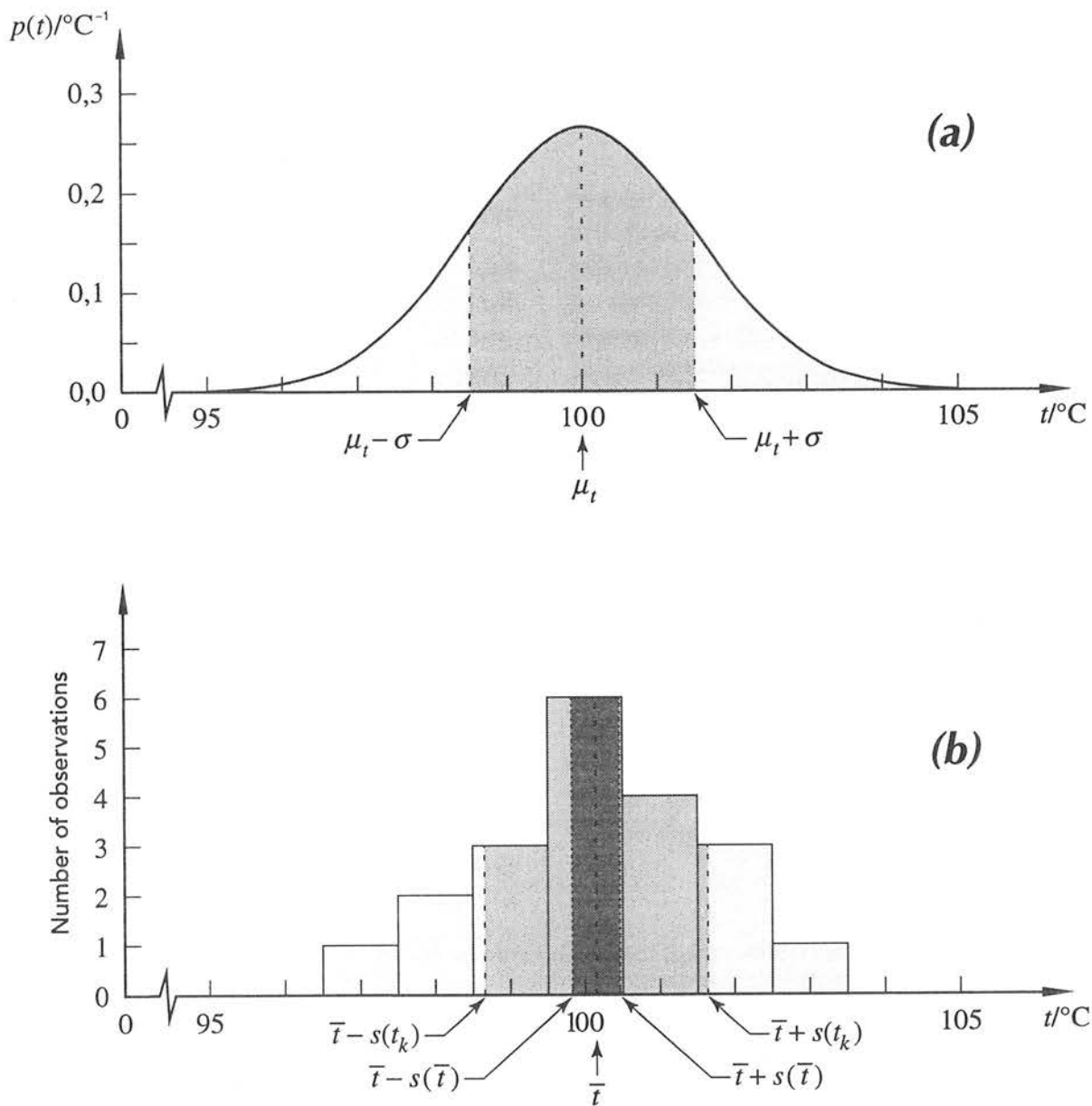


**Figure 1. Graphical illustration of evaluating the standard uncertainty of an input quantity from repeated observations**

is a temperature  $t$  and that its unknown distribution is a normal distribution with expectation  $\mu_t = 100\text{ }^{\circ}\text{C}$  and standard deviation  $\sigma = 1,5\text{ }^{\circ}\text{C}$ . Its probability density function (see C.2.14) is then

$$p(t) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left[-(t - \mu_t)^2 / 2\sigma^2\right]$$

NOTE – The definition of a probability density function  $p(z)$  requires that the relation  $\int p(z) dz = 1$  is satisfied.

**4.4.3** Figure 1b shows a histogram of  $n = 20$  repeated observations  $t_k$  of the temperature  $t$  that are assumed to have been taken randomly from the distribution of figure 1a. To obtain the histogram, the 20 observations or samples, whose values are given in table 1, are grouped into intervals  $1\text{ }^{\circ}\text{C}$  wide. (Preparation of a histogram is, of course, not required for the statistical analysis of the data.)

The arithmetic mean or average  $\bar{t}$  of the  $n = 20$  observations calculated according to equation (3) is  $\bar{t} = 100,145\text{ }^{\circ}\text{C} \approx 100,14\text{ }^{\circ}\text{C}$  and is assumed to be the best estimate of the expectation  $\mu_t$  of  $t$  based on the available data. The experimental standard deviation  $s(t_k)$  calculated from equation (4) is  $s(t_k) = 1,489\text{ }^{\circ}\text{C} \approx 1,49\text{ }^{\circ}\text{C}$ , and the experimental standard deviation of the mean  $s(\bar{t})$  calculated from equation (5), which is the standard uncertainty  $u(\bar{t})$  of the mean  $\bar{t}$ , is  $u(\bar{t}) = s(\bar{t}) = s(t_k)/\sqrt{20} = 0,333\text{ }^{\circ}\text{C} \approx 0,33\text{ }^{\circ}\text{C}$ . (For further calculations, it is likely that all of the digits would be retained.)

NOTE – Although the data in table 1 are not implausible considering the widespread use of high-resolution digital electronic thermometers, they are for illustrative purposes and should not necessarily be interpreted as describing a real measurement.

**4.4.4** Figure 2 represents the estimation of the value of an input quantity  $X_i$  and the evaluation of the uncertainty of that estimate from an *a priori* distribution of possible values of  $X_i$ , or probability distribution of  $X_i$ , based on all of the available information. For both cases shown, the input quantity is again assumed to be a temperature  $t$ .

**4.4.5** For the case illustrated in figure 2a, it is assumed that little information is available about the input quantity  $t$  and that all one can do is suppose that  $t$  is described by a symmetric, rectangular *a priori* probability distribution of lower bound  $a_- = 96\text{ }^{\circ}\text{C}$ , upper bound  $a_+ = 104\text{ }^{\circ}\text{C}$ , and thus half-width  $a = (a_+ - a_-)/2 = 4\text{ }^{\circ}\text{C}$  (see 4.3.7). The probability density function of  $t$  is then

$$p(t) = 1/2a, \quad a_- \leq t \leq a_+, \\ p(t) = 0, \quad \text{otherwise}$$

As indicated in 4.3.7, the best estimate of  $t$  is its expectation  $\mu_t = (a_+ + a_-)/2 = 100\text{ }^{\circ}\text{C}$ , which follows from C.3.1. The standard uncertainty of this estimate is  $u(\mu_t) = a/\sqrt{3} \approx 2,3\text{ }^{\circ}\text{C}$ , which follows from C.3.2 [see equation (7)].

**4.4.6** For the case illustrated in figure 2b, it is assumed that the available information concerning  $t$  is less limited

**Table 1** – Twenty repeated observations of the temperature  $t$  grouped in  $1\text{ }^{\circ}\text{C}$  intervals

| Interval $t_1 \leq t < t_2$ |                        | Temperature                                 |
|-----------------------------|------------------------|---------------------------------------------|
| $t_1/^{\circ}\text{C}$      | $t_2/^{\circ}\text{C}$ | $t/^{\circ}\text{C}$                        |
| 94,5                        | 95,5                   | —                                           |
| 95,5                        | 96,5                   | —                                           |
| 96,5                        | 97,5                   | 96,90                                       |
| 97,5                        | 98,5                   | 98,18; 98,25                                |
| 98,5                        | 99,5                   | 98,61; 99,03; 99,49                         |
| 99,5                        | 100,5                  | 99,56; 99,74; 99,89; 100,07; 100,33; 100,42 |
| 100,5                       | 101,5                  | 100,68; 100,95; 101,11; 101,20              |
| 101,5                       | 102,5                  | 101,57; 101,84; 102,36                      |
| 102,5                       | 103,5                  | 102,72                                      |
| 103,5                       | 104,5                  | —                                           |
| 104,5                       | 105,5                  | —                                           |

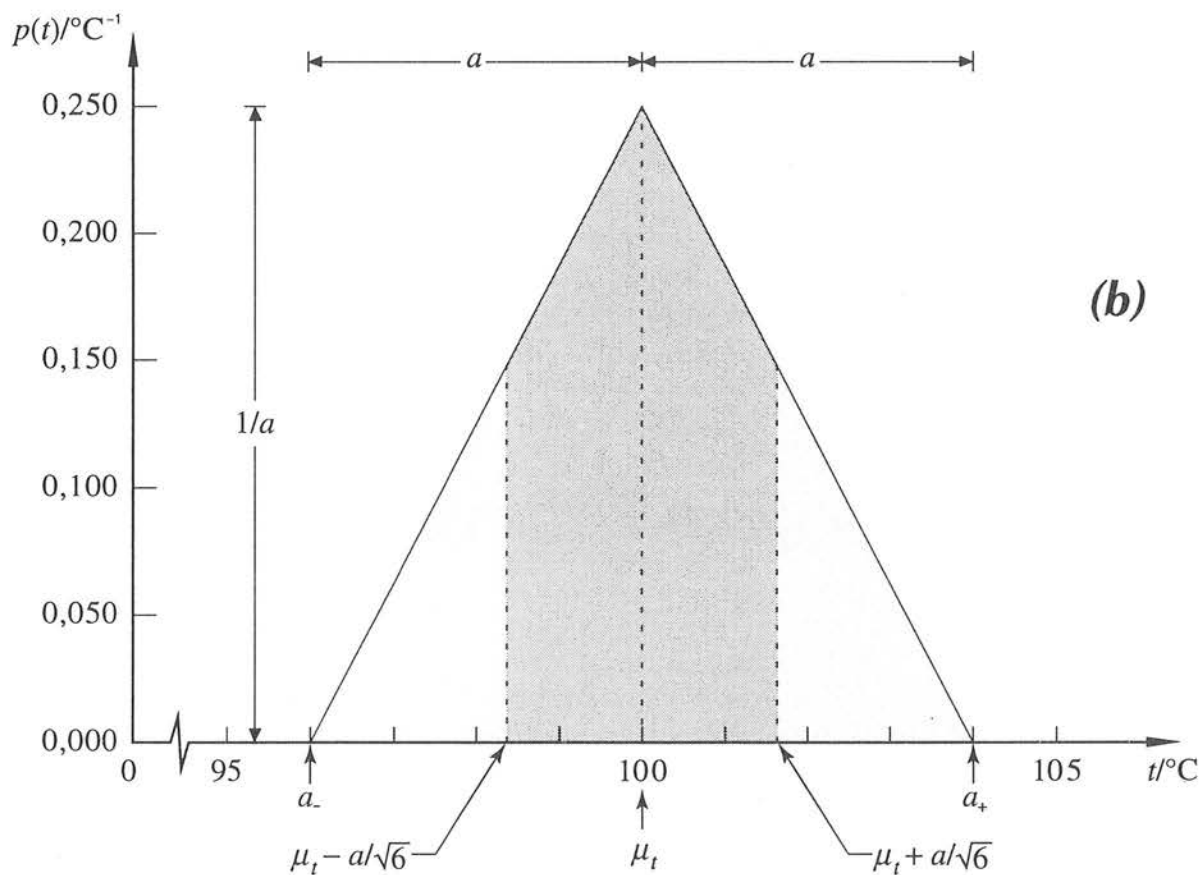
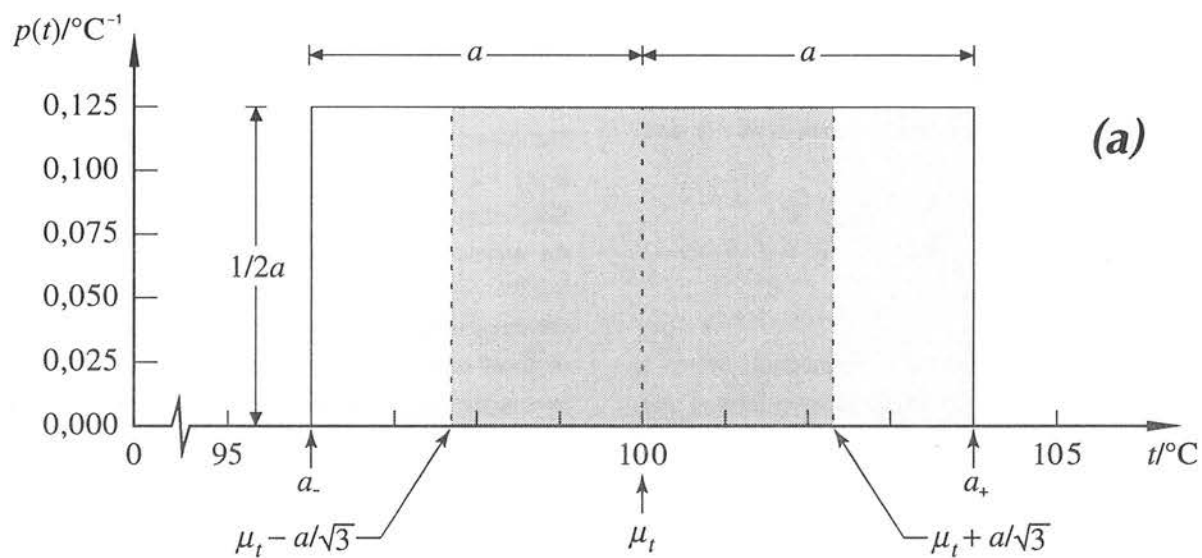


**Figure 2. Graphical illustration of evaluating the standard uncertainty of an input quantity from an a priori distribution**

and that  $t$  can be described by a symmetric, triangular *a priori* probability distribution of the same lower bound  $a_- = 96\text{ °C}$ , the same upper bound  $a_+ = 104\text{ °C}$ , and thus the same half-width  $a = (a_+ - a_-)/2 = 4\text{ °C}$  as in 4.4.5 (see 4.3.9). The probability density function of  $t$  is then

$$p(t) = (t - a_-)/a^2, \quad a_- \leq t \leq (a_+ + a_-)/2$$

$$p(t) = (a_+ - t)/a^2, \quad (a_+ + a_-)/2 \leq t \leq a_+$$

$$p(t) = 0, \quad \text{otherwise}$$

As indicated in 4.3.9, the expectation of  $t$  is  $\mu_t = (a_+ + a_-)/2 = 100\text{ °C}$ , which follows from C.3.1.

The standard uncertainty of this estimate is  $u(\mu_t) = a/\sqrt{6} \approx 1,6\text{ °C}$ , which follows from C.3.2 [see equation (9b)].

The above value,  $u(\mu_t) = 1,6\text{ °C}$ , may be compared with  $u(\mu_t) = 2,3\text{ °C}$  obtained in 4.4.5 from a rectangular distribution of the same  $8\text{ °C}$  width; with  $\sigma = 1,5\text{ °C}$  of the normal distribution of figure 1a whose  $-2,58\sigma$  to  $+2,58\sigma$  width, which encompasses 99 percent of the distribution, is nearly  $8\text{ °C}$ ; and with  $u(\bar{t}) = 0,33\text{ °C}$  obtained in 4.4.3 from 20 observations assumed to have been taken randomly from the same normal distribution.

## 5 Determining combined standard uncertainty

### 5.1 Uncorrelated input quantities

This subclause treats the case where all input quantities are **independent** (C.3.7). The case where two or more input quantities are related, that is, are interdependent or **correlated** (C.2.8), is discussed in 5.2.

**5.1.1** The standard uncertainty of  $y$ , where  $y$  is the estimate of the measurand  $Y$  and thus the result of the measurement, is obtained by appropriately combining the standard uncertainties of the input estimates  $x_1, x_2, \dots, x_N$  (see 4.1). This *combined standard uncertainty* of the estimate  $y$  is denoted by  $u_c(y)$ .

NOTE – For reasons similar to those given in the note to 4.3.1, the symbols  $u_c(y)$  and  $u_c^2(y)$  are used in all cases.

**5.1.2** The combined standard uncertainty  $u_c(y)$  is the positive square root of the combined variance  $u_c^2(y)$ , which is given by

$$u_c^2(y) = \sum_{i=1}^N \left[ \frac{\partial f}{\partial x_i} \right]^2 u^2(x_i) \quad \dots (10)$$

where  $f$  is the function given in equation (1). Each  $u(x_i)$  is a standard uncertainty evaluated as described in 4.2 (Type A evaluation) or as in 4.3 (Type B evaluation). The combined standard uncertainty  $u_c(y)$  is an estimated standard deviation and characterizes the dispersion of the values that could reasonably be attributed to the measurand  $Y$  (see 2.2.3).

Equation (10) and its counterpart for correlated input quantities, equation (13), both of which are based on a first-order Taylor series approximation of  $Y = f(X_1, X_2, \dots, X_N)$ , express what is termed in this *Guide* the *law of propagation of uncertainty* (see E.3.1 and E.3.2).

NOTE – When the nonlinearity of  $f$  is significant, higher-order terms in the Taylor series expansion must be included in the

expression for  $u_c^2(y)$ , equation (10). When the distribution of each  $X_i$  is symmetric about its mean, the most important terms of next highest order to be added to the terms of equation (10) are

$$\sum_{i=1}^N \sum_{j=1}^N \left[ \frac{1}{2} \left( \frac{\partial^2 f}{\partial x_i \partial x_j} \right)^2 + \frac{\partial f}{\partial x_i} \frac{\partial^3 f}{\partial x_i \partial x_j^2} \right] u^2(x_i) u^2(x_j)$$

See H.1 for an example of a situation where the contribution of higher-order terms to  $u_c^2(y)$  needs to be considered.

**5.1.3** The partial derivatives  $\partial f / \partial x_i$  are equal to  $\partial f / \partial X_i$  evaluated at  $X_i = x_i$  (see note 1 below). These derivatives, often called sensitivity coefficients, describe how the output estimate  $y$  varies with changes in the values of the input estimates  $x_1, x_2, \dots, x_N$ . In particular, the change in  $y$  produced by a small change  $\Delta x_i$  in input estimate  $x_i$  is given by  $(\Delta y)_i = (\partial f / \partial x_i)(\Delta x_i)$ . If this change is generated by the standard uncertainty of the estimate  $x_i$ , the corresponding variation in  $y$  is  $(\partial f / \partial x_i)u(x_i)$ . The combined variance  $u_c^2(y)$  can therefore be viewed as a sum of terms, each of which represents the estimated variance associated with the output estimate  $y$  generated by the estimated variance associated with each input estimate  $x_i$ . This suggests writing equation (10) as

$$u_c^2(y) = \sum_{i=1}^N [c_i u(x_i)]^2 = \sum_{i=1}^N u_i^2(y) \quad \dots (11a)$$

where

$$c_i \equiv \partial f / \partial x_i, \quad u_i(y) \equiv |c_i| u(x_i) \quad \dots (11b)$$

#### NOTES

1 Strictly speaking, the partial derivatives are  $\partial f / \partial x_i = \partial f / \partial X_i$  evaluated at the expectations of the  $X_i$ . However, in practice, the partial derivatives are estimated by

$$\frac{\partial f}{\partial x_i} = \frac{\partial f}{\partial X_i} \bigg|_{x_1, x_2, \dots, x_N}$$

2 The combined standard uncertainty  $u_c(y)$  may be calculated numerically by replacing  $c_i u(x_i)$  in equation (11a) with

$$Z_i = \frac{1}{2} [f(x_1, \dots, x_i + u(x_i), \dots, x_N) - f(x_1, \dots, x_i - u(x_i), \dots, x_N)]$$

That is,  $u_i(y)$  is evaluated numerically by calculating the change in  $y$  due to a change in  $x_i$  of  $+u(x_i)$  and of  $-u(x_i)$ . The value of  $u_i(y)$  may then be taken as  $|Z_i|$  and the value of the corresponding sensitivity coefficient  $c_i$  as  $Z_i/u(x_i)$ .

EXAMPLE – For the example of 4.1.1, using the same symbol for both the quantity and its estimate for simplicity of notation,

$$\begin{aligned} c_1 &\equiv \partial P / \partial V = 2V/R_0[1 + \alpha(t - t_0)] = 2P/V \\ c_2 &\equiv \partial P / \partial R_0 = -V^2/R_0^2[1 + \alpha(t - t_0)] = -P/R_0 \\ c_3 &\equiv \partial P / \partial \alpha = -V^2(t - t_0)/R_0[1 + \alpha(t - t_0)]^2 \\ &\quad = -P(t - t_0)/[1 + \alpha(t - t_0)] \\ c_4 &\equiv \partial P / \partial t = -V^2\alpha/R_0[1 + \alpha(t - t_0)]^2 \\ &\quad = -P\alpha/[1 + \alpha(t - t_0)] \end{aligned}$$

and

$$\begin{aligned} u^2(P) &= \left[ \frac{\partial P}{\partial V} \right]^2 u^2(V) + \left[ \frac{\partial P}{\partial R_0} \right]^2 u^2(R_0) \\ &\quad + \left[ \frac{\partial P}{\partial \alpha} \right]^2 u^2(\alpha) + \left[ \frac{\partial P}{\partial t} \right]^2 u^2(t) \\ &= [c_1 u(V)]^2 + [c_2 u(R_0)]^2 + [c_3 u(\alpha)]^2 + [c_4 u(t)]^2 \\ &= u_1^2(P) + u_2^2(P) + u_3^2(P) + u_4^2(P) \end{aligned}$$

**5.1.4** Instead of being calculated from the function  $f$ , sensitivity coefficients  $\partial f / \partial x_i$  are sometimes determined experimentally: one measures the change in  $Y$  produced by a change in a particular  $X_i$  while holding the remaining input quantities constant. In this case, the knowledge of the function  $f$  (or a portion of it when only several sensitivity coefficients are so determined) is accordingly reduced to an empirical first-order Taylor series expansion based on the measured sensitivity coefficients.

**5.1.5** If equation (1) for the measurand  $Y$  is expanded about nominal values  $X_{i,0}$  of the input quantities  $X_i$ , then, to first order (which is usually an adequate approximation),  $Y = Y_0 + c_1 \delta_1 + c_2 \delta_2 + \dots + c_N \delta_N$ , where  $Y_0 = f(X_{1,0}, X_{2,0}, \dots, X_{N,0})$ ,  $c_i = (\partial f / \partial X_i)$  evaluated at  $X_i = X_{i,0}$ , and  $\delta_i = X_i - X_{i,0}$ . Thus, for the purposes of an analysis of uncertainty, a measurand is usually approximated by a linear function of its variables by transforming its input quantities from  $X_i$  to  $\delta_i$  (see E.3.1).

EXAMPLE – From example 2 of 4.3.7, the estimate of the value of the measurand  $V$  is  $V = \bar{V} + \Delta \bar{V}$ , where  $\bar{V} = 0,928\,571\text{ V}$ ,  $u(\bar{V}) = 12\text{ }\mu\text{V}$ , the additive correction  $\Delta \bar{V} = 0$ , and  $u(\Delta \bar{V}) = 8,7\text{ }\mu\text{V}$ . Since  $\partial V / \partial \bar{V} = 1$  and  $\partial V / \partial (\Delta \bar{V}) = 1$ , the combined variance associated with  $V$  is given by

$$\begin{aligned} u_c^2(V) &= u^2(\bar{V}) + u^2(\Delta \bar{V}) = (12\text{ }\mu\text{V})^2 + (8,7\text{ }\mu\text{V})^2 \\ &= 219 \times 10^{-12}\text{ V}^2 \end{aligned}$$

and the combined standard uncertainty is  $u_c(V) = 15\text{ }\mu\text{V}$ , which corresponds to a relative combined standard uncertainty  $u_c(V)/V$  of  $16 \times 10^{-6}$  (see 5.1.6). This is an example of the case where the measurand is already a linear function of the quantities on which it depends, with coefficients  $c_i = +1$ . It follows from equation (10) that if  $Y = c_1 X_1 + c_2 X_2 + \dots + c_N X_N$  and if the constants  $c_i = +1$  or  $-1$ , then  $u_c^2(y) = \sum_{i=1}^N u^2(x_i)$ .

**5.1.6** If  $Y$  is of the form  $Y = c X_1^{p_1} X_2^{p_2} \dots X_N^{p_N}$  and the exponents  $p_i$  are known positive or negative numbers having negligible uncertainties, the combined variance, equation (10), can be expressed as

$$[u_c(y)/y]^2 = \sum_{i=1}^N [p_i u(x_i)/x_i]^2 \quad \dots (12)$$

This is of the same form as equation (11a) but with the combined variance  $u_c^2(y)$  expressed as a *relative combined variance*  $[u_c(y)/y]^2$  and the estimated variance  $u^2(x_i)$  associated with each input estimate expressed as an estimated *relative variance*  $[u(x_i)/x_i]^2$ . [The *relative combined standard uncertainty* is  $u_c(y)/|y|$  and the *relative standard uncertainty* of each input estimate is  $u(x_i)/|x_i|$ ,  $|y| \neq 0$  and  $|x_i| \neq 0$ .]

#### NOTES

1 When  $Y$  has this form, its transformation to a linear function of variables (see 5.1.5) is readily achieved by setting  $X_i = X_{i,0}(1 + \delta_i)$ , for then the following approximate relation results:  $(Y - Y_0)/Y_0 = \sum_{i=1}^N p_i \delta_i$ . On the other hand, the logarithmic transformation  $Z = \ln Y$  and  $W_i = \ln X_i$  leads to an exact linearization in terms of the new variables:  $Z = \ln c + \sum_{i=1}^N p_i W_i$ .

2 If each  $p_i$  is either  $+1$  or  $-1$ , equation (12) becomes  $[u_c(y)/y]^2 = \sum_{i=1}^N [u(x_i)/x_i]^2$ , which shows that for this special case the relative combined variance associated with the estimate  $y$  is simply equal to the sum of the estimated relative variances associated with the input estimates  $x_i$ .

## 5.2 Correlated input quantities

**5.2.1** Equation (10) and those derived from it such as equations (11) and (12) are valid only if the input

quantities  $X_i$  are independent or uncorrelated (the random variables, not the physical quantities that are assumed to be invariants – see 4.1.1, note 1). If some of the  $X_i$  are significantly correlated, the correlations must be taken into account.

5.2.2 When the input quantities are correlated, the appropriate expression for the combined variance  $u_c^2(y)$  associated with the result of a measurement is

$$\begin{aligned} u_c^2(y) &= \sum_{i=1}^N \sum_{j=1}^N \frac{\partial f}{\partial x_i} \frac{\partial f}{\partial x_j} u(x_i, x_j) \\ &= \sum_{i=1}^N \left[ \frac{\partial f}{\partial x_i} \right]^2 u^2(x_i) \quad \dots (13) \\ &\quad + 2 \sum_{i=1}^{N-1} \sum_{j=i+1}^N \frac{\partial f}{\partial x_i} \frac{\partial f}{\partial x_j} u(x_i, x_j) \end{aligned}$$

where  $x_i$  and  $x_j$  are the estimates of  $X_i$  and  $X_j$  and  $u(x_i, x_j) = u(x_j, x_i)$  is the estimated covariance associated with  $x_i$  and  $x_j$ . The degree of correlation between  $x_i$  and  $x_j$  is characterized by the estimated **correlation coefficient** (C.3.6)

$$r(x_i, x_j) = \frac{u(x_i, x_j)}{u(x_i) u(x_j)} \quad \dots (14)$$

where  $r(x_i, x_j) = r(x_j, x_i)$ , and  $-1 \leq r(x_i, x_j) \leq +1$ . If the estimates  $x_i$  and  $x_j$  are independent,  $r(x_i, x_j) = 0$ , and a change in one does not imply an expected change in the other. (See C.2.8, C.3.6, and C.3.7 for further discussion.)

In terms of correlation coefficients, which are more readily interpreted than covariances, the covariance term of equation (13) may be written as

$$2 \sum_{i=1}^{N-1} \sum_{j=i+1}^N \frac{\partial f}{\partial x_i} \frac{\partial f}{\partial x_j} u(x_i) u(x_j) r(x_i, x_j) \quad \dots (15)$$

Equation (13) then becomes, with the aid of equation (11b),

$$\begin{aligned} u_c^2(y) &= \sum_{i=1}^N c_i^2 u^2(x_i) \\ &\quad + 2 \sum_{i=1}^{N-1} \sum_{j=i+1}^N c_i c_j u(x_i) u(x_j) r(x_i, x_j) \quad \dots (16) \end{aligned}$$

#### NOTES

1 For the very special case where *all* of the input estimates are correlated with correlation coefficients  $r(x_i, x_j) = +1$ , equation

(16) reduces to

$$u_c^2(y) = \left[ \sum_{i=1}^N c_i u(x_i) \right]^2 = \left[ \sum_{i=1}^N \frac{\partial f}{\partial x_i} u(x_i) \right]^2$$

The combined standard uncertainty  $u_c(y)$  is thus simply a *linear sum* of terms representing the variation of the output estimate  $y$  generated by the standard uncertainty of each input estimate  $x_i$  (see 5.1.3). [This linear sum should not be confused with the general law of error propagation although it has a similar form; standard uncertainties are not errors (see E.3.2).]

**EXAMPLE** – Ten resistors, each of nominal resistance  $R_i = 1000 \, \Omega$ , are calibrated with a negligible uncertainty of comparison in terms of the same  $1000 \, \Omega$  standard resistor  $R_S$  characterized by a standard uncertainty  $u(R_S) = 100 \, \text{m}\Omega$  as given in its calibration certificate. The resistors are connected in series with wires having negligible resistance in order to obtain a reference resistance  $R_{\text{ref}}$  of nominal value  $10 \, \text{k}\Omega$ . Thus  $R_{\text{ref}} = f(R_i) = \sum_{i=1}^{10} R_i$ . Since  $r(x_i, x_j) = r(R_i, R_j) = +1$  for each resistor pair (see F.1.2.3, example 2), the equation of this note applies. Since for each resistor  $\partial f / \partial x_i = \partial R_{\text{ref}} / \partial R_i = 1$ , and  $u(x_i) = u(R_i) = u(R_S)$  (see F.1.2.3, example 2), that equation yields for the combined standard uncertainty of  $R_{\text{ref}}$ ,  $u_c(R_{\text{ref}}) = \sum_{i=1}^{10} u(R_S) = 10 \times (100 \, \text{m}\Omega) = 1 \, \Omega$ . The result  $u_c(R_{\text{ref}}) = [\sum_{i=1}^{10} u^2(R_S)]^{1/2} = 0,32 \, \Omega$  obtained from equation (10) is incorrect because it does not take into account that all of the calibrated values of the ten resistors are correlated.

2 The estimated variances  $u^2(x_i)$  and estimated covariances  $u(x_i, x_j)$  may be considered as the elements of a covariance matrix with elements  $u_{ij}$ . The diagonal elements  $u_{ii}$  of the matrix are the variances  $u^2(x_i)$ , while the off-diagonal elements  $u_{ij}$  ( $i \neq j$ ) are the covariances  $u(x_i, x_j) = u(x_j, x_i)$ . If two input estimates are uncorrelated, their associated covariance and the corresponding elements  $u_{ij}$  and  $u_{ji}$  of the covariance matrix are 0. If the input estimates are all uncorrelated, all of the off-diagonal elements are zero and the covariance matrix is diagonal. (See also C.3.5.)

3 For the purposes of numerical evaluation, equation (16) may be written as

$$u_c^2(y) = \sum_{i=1}^N \sum_{j=1}^N Z_i Z_j r(x_i, x_j)$$

where  $Z_i$  is given in 5.1.3, note 2.

4 If the  $X_i$  of the special form considered in 5.1.6 are correlated, then the terms

$$2 \sum_{i=1}^{N-1} \sum_{j=i+1}^N [p_i u(x_i) / x_i] [p_j u(x_j) / x_j] r(x_i, x_j)$$

must be added to the right-hand side of equation (12).

5.2.3 Consider two arithmetic means  $\bar{q}$  and  $\bar{r}$  that

estimate the expectations  $\mu_q$  and  $\mu_r$  of two randomly varying quantities  $q$  and  $r$ , and let  $\bar{q}$  and  $\bar{r}$  be calculated from  $n$  independent pairs of simultaneous observations of  $q$  and  $r$  made under the same conditions of measurement (see B.2.15). Then the covariance (see C.3.4) of  $\bar{q}$  and  $\bar{r}$  is estimated by

$$s(\bar{q}, \bar{r}) = \frac{1}{n(n-1)} \sum_{k=1}^n (q_k - \bar{q})(r_k - \bar{r}) \quad \dots (17)$$

where  $q_k$  and  $r_k$  are the individual observations of the quantities  $q$  and  $r$ , and  $\bar{q}$  and  $\bar{r}$  are calculated from the observations according to equation (3). If in fact the observations are uncorrelated, the calculated covariance is expected to be near 0.

Thus the estimated covariance of two correlated input quantities  $X_i$  and  $X_j$  that are estimated by the means  $\bar{X}_i$  and  $\bar{X}_j$  determined from independent pairs of repeated simultaneous observations is given by  $u(x_i, x_j) = s(\bar{X}_i, \bar{X}_j)$ , with  $s(\bar{X}_i, \bar{X}_j)$  calculated according to equation (17). This application of equation (17) is a Type A evaluation of covariance. The estimated correlation coefficient of  $\bar{X}_i$  and  $\bar{X}_j$  is obtained from equation (14):  $r(x_i, x_j) = r(\bar{X}_i, \bar{X}_j) = s(\bar{X}_i, \bar{X}_j)/s(\bar{X}_i)s(\bar{X}_j)$ .

NOTE – Examples where it is necessary to use covariances as calculated from equation (17) are given in H.2 and H.4.

**5.2.4** There may be significant correlation between two input quantities if the same measuring instrument, physical measurement standard, or reference datum having a significant standard uncertainty is used in their determination. For example, if a certain thermometer is

used to determine a temperature correction required in the estimation of the value of input quantity  $X_i$ , and the same thermometer is used to determine a similar temperature correction required in the estimation of input quantity  $X_j$ , the two input quantities could be significantly correlated. However, if  $X_i$  and  $X_j$  in this example are redefined to be the uncorrected quantities and the quantities that define the calibration curve for the thermometer are included as additional input quantities with independent standard uncertainties, the correlation between  $X_i$  and  $X_j$  is removed. (See F.1.2.3 and F.1.2.4 for further discussion.)

**5.2.5** Correlations between input quantities cannot be ignored if present and significant. The associated covariances should be evaluated experimentally if feasible by varying the correlated input quantities (see C.3.6, note 3), or by using the pool of available information on the correlated variability of the quantities in question (Type B evaluation of covariance). Insight based on experience and general knowledge (see 4.3.1 and 4.3.2) is especially required when estimating the degree of correlation between input quantities arising from the effects of common influences, such as ambient temperature, barometric pressure, and humidity. Fortunately, in many cases, the effects of such influences have negligible interdependence and the affected input quantities can be assumed to be uncorrelated. However, if they cannot be assumed to be uncorrelated, the correlations themselves can be avoided if the common influences are introduced as additional independent input quantities as indicated in 5.2.4.

## 6 Determining expanded uncertainty

### 6.1 Introduction

$$U = k u_c(y) \quad \dots (18)$$

**6.1.1** Recommendation INC-1 (1980) of the Working Group on the Statement of Uncertainties on which this *Guide* is based (see the Introduction), and Recommendations 1 (CI-1981) and 1 (CI-1986) of the CIPM approving and reaffirming INC-1 (1980) (see A.2 and A.3), advocate the use of the combined standard uncertainty  $u_c(y)$  as the parameter for expressing quantitatively the uncertainty of the result of a measurement. Indeed, in the second of its recommendations, the CIPM has requested that what is now termed combined standard uncertainty  $u_c(y)$  be used “by all participants in giving the results of all international comparisons or other work done under the auspices of the CIPM and Comités Consultatifs.”

**6.1.2** Although  $u_c(y)$  can be universally used to express the uncertainty of a measurement result, in some commercial, industrial, and regulatory applications, and when health and safety are concerned, it is often necessary to give a measure of uncertainty that defines an interval about the measurement result that may be expected to encompass a large fraction of the distribution of values that could reasonably be attributed to the measurand. The existence of this requirement was recognized by the Working Group and led to paragraph 5 of Recommendation INC-1 (1980). It is also reflected in Recommendation 1 (CI-1986) of the CIPM.

### 6.2 Expanded uncertainty

**6.2.1** The additional measure of uncertainty that meets the requirement of providing an interval of the kind indicated in 6.1.2 is termed *expanded uncertainty* and is denoted by  $U$ . The expanded uncertainty  $U$  is obtained by multiplying the combined standard uncertainty  $u_c(y)$  by a *coverage factor*  $k$ :

The result of a measurement is then conveniently expressed as  $Y = y \pm U$ , which is interpreted to mean that the best estimate of the value attributable to the measurand  $Y$  is  $y$ , and that  $y - U$  to  $y + U$  is an interval that may be expected to encompass a large fraction of the distribution of values that could reasonably be attributed to  $Y$ . Such an interval is also expressed as  $y - U \leq Y \leq y + U$ .

**6.2.2** The terms **confidence interval** (C.2.27, C.2.28) and **confidence level** (C.2.29) have specific definitions in statistics and are only applicable to the interval defined by  $U$  when certain conditions are met, including that all components of uncertainty that contribute to  $u_c(y)$  be obtained from Type A evaluations. Thus, in this *Guide*, the word “confidence” is not used to modify the word “interval” when referring to the interval defined by  $U$ ; and the term “confidence level” is not used in connection with that interval but rather the term “level of confidence.” More specifically,  $U$  is interpreted as defining an interval about the measurement result that encompasses a large fraction  $p$  of the probability distribution characterized by that result and its combined standard uncertainty, and  $p$  is the *coverage probability* or *level of confidence* of the interval.

**6.2.3** Whenever practicable, the level of confidence  $p$  associated with the interval defined by  $U$  should be estimated and stated. It should be recognized that multiplying  $u_c(y)$  by a constant provides no new information but presents the previously available information in a different form. However, it should also be recognized that in most cases the level of confidence  $p$  (especially for values of  $p$  near 1) is rather uncertain, not only because of limited knowledge of the probability distribution characterized by  $y$  and  $u_c(y)$  (particularly in the extreme portions), but also because of the uncertainty

of  $u_c(y)$  itself (see note 2 to 2.3.5, 6.3.2, and annex G, especially G.6.6).

NOTE – For preferred ways of stating the result of a measurement when the measure of uncertainty is  $u_c(y)$  and when it is  $U$ , see 7.2.2 and 7.2.4, respectively.

### 6.3 Choosing a coverage factor

**6.3.1** The value of the coverage factor  $k$  is chosen on the basis of the level of confidence required of the interval  $y - U$  to  $y + U$ . In general,  $k$  will be in the range 2 to 3. However, for special applications  $k$  may be outside this range. Extensive experience with and full knowledge of the uses to which a measurement result will be put can facilitate the selection of a proper value of  $k$ .

NOTE – Occasionally, one may find that a known correction  $b$  for a systematic effect has not been applied to the reported result of a measurement, but instead an attempt is made to take the effect into account by enlarging the “uncertainty” assigned to the result. This should be avoided; only in very special circumstances should corrections for known significant systematic effects not be applied to the result of a measurement (see F.2.4.5 for a specific case and how to treat it). Evaluating the uncertainty of a measurement result should not be confused with assigning a safety limit to some quantity.

**6.3.2** Ideally, one would like to be able to choose a specific value of the coverage factor  $k$  that would provide an interval  $Y = y \pm U = y \pm k u_c(y)$  corresponding to a particular level of confidence  $p$ , such as 95 or 99 percent;

equivalently, for a given value of  $k$ , one would like to be able to state unequivocally the level of confidence associated with that interval. However, this is not easy to do in practice because it requires extensive knowledge of the probability distribution characterized by the measurement result  $y$  and its combined standard uncertainty  $u_c(y)$ . Although these parameters are of critical importance, they are by themselves insufficient for the purpose of establishing intervals having exactly known levels of confidence.

**6.3.3** Recommendation INC-1 (1980) does not specify how the relation between  $k$  and  $p$  should be established. This problem is discussed in annex G, and a preferred method for its approximate solution is presented in G.4 and summarized in G.6.4. However, a simpler approach, discussed in G.6.6, is often adequate in measurement situations where the probability distribution characterized by  $y$  and  $u_c(y)$  is approximately normal and the effective degrees of freedom of  $u_c(y)$  is of significant size. When this is the case, which frequently occurs in practice, one can assume that taking  $k = 2$  produces an interval having a level of confidence of approximately 95 percent, and that taking  $k = 3$  produces an interval having a level of confidence of approximately 99 percent.

NOTE – A method for estimating the effective degrees of freedom of  $u_c(y)$  is given in G.4. Table G.2 of annex G can then be used to help decide if this solution is appropriate for a particular measurement (see G.6.6).

## 7 Reporting uncertainty

### 7.1 General guidance

**7.1.1** In general, as one moves up the measurement hierarchy, more details are required on how a measurement result and its uncertainty were obtained. Nevertheless, at any level of this hierarchy, including commercial and regulatory activities in the marketplace, engineering work in industry, lower-echelon calibration facilities, industrial research and development, academic research, industrial primary standards and calibration laboratories, and the national standards laboratories and the BIPM, all of the information necessary for the re-evaluation of the measurement should be available to others who may have need of it. The primary difference is that at the lower levels of the hierarchical chain, more of the necessary information may be made available in the form of published calibration and test system reports, test specifications, calibration and test certificates, instruction manuals, international standards, national standards, and local regulations.

**7.1.2** When the details of a measurement, including how the uncertainty of the result was evaluated, are provided by referring to published documents, as is often the case when calibration results are reported on a certificate, it is imperative that these publications be kept up-to-date so that they are consistent with the measurement procedure actually in use.

**7.1.3** Numerous measurements are made every day in industry and commerce without any explicit report of uncertainty. However, many are performed with instruments subject to periodic calibration or legal inspection. If the instruments are known to be in conformance with their specifications or with the existing normative documents that apply, the uncertainties of their indications may be inferred from these specifications or from these normative documents.

**7.1.4** Although in practice the amount of information necessary to document a measurement result depends on its intended use, the basic principle of what is required remains unchanged: when reporting the result of a measurement and its uncertainty, it is preferable to err on the side of providing too much information rather than too little. For example, one should

- a) describe clearly the methods used to calculate the measurement result and its uncertainty from the experimental observations and input data;
- b) list all uncertainty components and document fully how they were evaluated;
- c) present the data analysis in such a way that each of its important steps can be readily followed and the calculation of the reported result can be independently repeated if necessary;
- d) give all corrections and constants used in the analysis and their sources.

A test of the foregoing list is to ask oneself "Have I provided enough information in a sufficiently clear manner that my result can be updated in the future if new information or data become available?"

### 7.2 Specific guidance

**7.2.1** When reporting the result of a measurement, and when the measure of uncertainty is the combined standard uncertainty  $u_c(y)$ , one should

- a) give a full description of how the measurand  $Y$  is defined;
- b) give the estimate  $y$  of the measurand  $Y$  and its combined standard uncertainty  $u_c(y)$ ; the units of  $y$  and  $u_c(y)$  should always be given;
- c) include the relative combined standard uncertainty  $u_c(y)/|y|$ ,  $|y| \neq 0$ , when appropriate;