GUIDE 98-3

Uncertainty of measurement —

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

Incertitude de mesure —

Partie 3: Guide pour l'expression de l'incertitude de mesure (GUM:1995)

First edition 2008

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

Published in Switzerland

Contents

Page

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***.

* **Footnote to the 2008 version:**

In 2005, the International Laboratory Accreditation Cooperation (ILAC) officially joined the seven founding international organizations.

** **Footnote to the 2008 version:**

The names of these three organizations have changed since 1995. They are now:
IFCC: International Federation for Clinical Chemistry and Laboratory Medicine
IUPAC: International Organization for Pure and Applied Chemistry
IUPAP: International Organization for Pure and Applied Physics.

*** **Footnote to the 2008 version:**

Links to the addresses of the eight organizations presently involved in the JCGM (Joint Committee for Guides in Metrology) are given on http://www.bipm.org/en/committees/jc/jcgm.

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].[1)] 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.

This first edition of ISO/IEC Guide 98-3 cancels and replaces the *Guide to the Expression of Uncertainty in Measurement (GUM)*, BIPM, IEC, IFCC, ISO, IUPAC, IUPAP, OIML, 1993, corrected and reprinted in 1995.

---

1) See the Bibliography.

* **Footnote to the 2008 version:**

In producing this 2008 version of the GUM, necessary corrections only to the printed 1995 version have been introduced by JCGM/WG 1. These corrections occur in Subclauses 4.2.2, 4.2.4, 5.1.2, B.2.17, C.3.2, C.3.4, E.4.3, H.4.3, H.5.2.5 and H.6.2.

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 (Bibliography). 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 v_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.

Uncertainty of measurement —

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

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).

* **Footnote to the 2008 version:**

Several derivative general and specific applications documents have been published. Non-exhaustive compilations listing these documents can be found on http://www.bipm.org/en/committees/jc/jcgm/wg1_bibliography.html. In addition, up-to-date listings of documents that cite the *Guide to the expression of uncertainty in measurement* can be found by using the full-text search options on http://www.iso.org/ and http://www.iec.ch/.

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 VIM [6] (VIM:1993, definition 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

NOTE 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.

NOTE 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.

NOTE 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

* **Footnote to the 2008 version:**
The third edition of the vocabulary was published in 2007, under the title ISO/IEC Guide 99, *International vocabulary of metrology — Basic and general concepts and associated terms (VIM)*.

. 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:1984, definition 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

NOTE 1 The fraction may be viewed as the coverage probability or level of confidence of the interval.

NOTE 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.

NOTE 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).

* **Footnote to the 2008 version:**

According to Resolution 10 of the 22nd CGPM (2003) "... the symbol for the decimal marker shall be either the point on the line or the comma on the line...". The JCGM has decided to adopt, in its documents in English, the point on the line. However, in this document, the decimal comma has been retained for consistency with the 1995 printed version.

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.

NOTE 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.

NOTE 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.

NOTE 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.

NOTE 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 $2 \cdot 10^{-8}$ when V_S is reported in terms of the conventional value, but $u_c(V_S)/V_S$ is $4 \cdot 10^{-7}$ 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 $X_1, X_2, ..., X_N$ through a functional relationship f:

$$Y = f(X_1, X_2, ..., X_N) \tag{1}$$

NOTE 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).

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

NOTE 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 α, the power P (the measurand) dissipated by the resistor at the temperature t depends on V, R_0, α, 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* $X_1, X_2, ..., X_N$ 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 $X_1, X_2, ..., X_N$ 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* $x_1, x_2, ..., x_N$ for the values of the N quantities $X_1, X_2, ..., X_N$. Thus the *output estimate* y, which is the result of the measurement, is given by

$$y = f(x_1, x_2, ..., x_N) \tag{2}$$

NOTE In some cases, the estimate y may be obtained from

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

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 $y = f(\bar{X}_1, \bar{X}_2, ..., \bar{X}_N)$, where

$$\bar{X}_i = \frac{1}{n}\sum_{k=1}^{n} X_{i,k}$$

is the arithmetic mean of the individual observations $X_{i,k}$, may be preferable when f is a nonlinear function of the input quantities $X_1, X_2, ..., X_N$, 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 μ_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:

$$\bar{q} = \frac{1}{n}\sum_{k=1}^{n} q_k \tag{3}$$

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 σ^2 of the probability distribution of q, is given by

$$s^2(q_k) = \frac{1}{n-1}\sum_{j=1}^{n}(q_j - \bar{q})^2 \tag{4}$$

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

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

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 μ_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.

NOTE 1 The number of observations n should be large enough to ensure that $\bar{q}$ provides a reliable estimate of the expectation μ_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).

NOTE 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 $s^2(q_k)/n$ 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) v_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 µg at the three standard deviation level". The standard uncertainty of the mass standard is then simply $u(m_S) = (240\ \mu g)/3 = 80\ \mu g$. This corresponds to a relative standard uncertainty $u(m_S)/m_S$ of $80 \cdot 10^{-9}$ (see 5.1.6). The estimated variance is $u^2(m_S) = (80\ \mu g)^2 = 6{,}4 \cdot 10^{-9}\ g^2$.

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 . ± 129 µ. at 23 °C and that "the quoted uncertainty of 129 µ. defines an interval having a level of confidence of 99 percent". The standard uncertainty of the resistor may be taken as $u(R_S) = (129\ \mu.)/2{,}58 = 50\ \mu.$, which corresponds to a relative standard uncertainty $u(R_S)/R_S$ of $5{,}0 \cdot 10^{-6}$ (see 5.1.6). The estimated variance is $u^2(R_S) = (50\ \mu.)^2 = 2{,}5 \cdot 10^{-9}\ .^2$

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 $\propto$ and standard deviation σ the interval $\propto \pm\ \sigma/1{,}48$ 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 ± 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 $u(l) = 1{,}48 \cdot 0{,}04\ mm \quad 0{,}06\ mm$ and the estimated variance is $u^2(l) = (1{,}48 \cdot 0{,}04\ mm)^2 = 3{,}5 \cdot 10^{-3}\ mm^2$.

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 $\propto$ and standard deviation σ the interval $\propto \pm\ \sigma$ 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 0,967 42 corresponding to probability $p = 2/3$, that is, if one were to write $u(x_i) = a/0{,}967\ 42 = 1{,}033a$.

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 2 a). 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

$$u^2(x_i) = (a_+ - a_-)^2/12 \tag{6}$$

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

$$u^2(x_i) = a^2/3 \tag{7}$$

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.

EXAMPLE 1 A handbook gives the value of the coefficient of linear thermal expansion of pure copper at 20 °C, α_{20}(Cu), as 16,52 · 10^{-6} °C^{-1} and simply states that "the error in this value should not exceed 0,40 · 10^{-6} °C^{-1}". Based on this limited information, it is not unreasonable to assume that the value of α_{20}(Cu) lies with equal probability in the interval 16,12 · 10^{-6} °C^{-1} to 16,92 · 10^{-6} °C^{-1}, and that it is very unlikely that α_{20}(Cu) lies outside this interval. The variance of this symmetric rectangular distribution of possible values of α_{20}(Cu) of half-width $a = 0{,}40 \cdot 10^{-6}$ °C^{-1} is then, from Equation (7), $u^2(\alpha_{20}) = (0{,}40 \cdot 10^{-6}\ °C^{-1})^2/3 = 53{,}3 \cdot 10^{-15}\ °C^{-2}$, and the standard uncertainty is $u(\alpha_{20}) = (0{,}40 \cdot 10^{-6}\ °C^{-1})/\sqrt{3} = 0{,}23 \cdot 10^{-6}\ °C^{-1}$.

EXAMPLE 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 · 10^{-6} times the reading plus 2 · 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 $\bar{V} = 0{,}928\ 571$ V with a Type A standard uncertainty $u(\bar{V}) = 12$ µV. 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 $\bar{V}$, . $\bar{V}$, 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 . $\bar{V}$ is then $a = (14 \cdot 10^{-6}) \cdot (0{,}928\ 571\ V) + (2 \cdot 10^{-6}) \cdot (1\ V) = 15$ µV, and from Equation (7), $u^2(.\ \bar{V}) = 75\ \mu V^2$ and $u(.\ \bar{V}) = 8{,}7$ µV. The estimate of the value of the measurand V, for simplicity denoted by the same symbol V, is given by $V = \bar{V} + .\ \bar{V} = 0{,}928\ 571$ V. One can obtain the combined standard uncertainty of this estimate by combining the 12 µV Type A standard uncertainty of $\bar{V}$ with the 8,7 µV Type B standard uncertainty of . $\bar{V}$. 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 $b_- > b_+$. 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

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

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 $\alpha_{20}(\mathrm{Cu}) = 16{,}52 \cdot 10^{-6}\ ^\circ\mathrm{C}^{-1}$ and it is stated that "the smallest possible value is $16{,}40 \cdot 10^{-6}\ ^\circ\mathrm{C}^{-1}$ and the largest possible value is $16{,}92 \cdot 10^{-6}\ ^\circ\mathrm{C}^{-1}$", then $b_- = 0{,}12 \cdot 10^{-6}\ ^\circ\mathrm{C}^{-1}$, $b_+ = 0{,}40 \cdot 10^{-6}\ ^\circ\mathrm{C}^{-1}$, and, from Equation (8), $u(\alpha_{20}) = 0{,}15 \cdot 10^{-6}\ ^\circ\mathrm{C}^{-1}$.

NOTE 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 $b_-' = b_+' = (b_+ + b_-)/2 = (a_+ - a_-)/2 = a$.

NOTE 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 $u^2(x_i) = b_+ b_- - (b_+ - b_-)/\lambda$; 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 $0 \leqslant \beta \leqslant 1$. As $\beta \to 1$, 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 2 b)]. 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

$$u^2(x_i) = a^2\left(1 + \beta^2\right)/6 \tag{9a}$$

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

$$u^2(x_i) = a^2/6 \tag{9b}$$

NOTE 1 For a normal distribution with expectation μ and standard deviation σ, 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.

NOTE 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 1 a), it is assumed that the input quantity X_i is a temperature t and that its unknown distribution is a normal distribution with expectation $\propto_t = 100\ °C$ and standard deviation $\sigma = 1{,}5\ °C$. Its probability density function (see C.2.14) is then

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

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

4.4.3 Figure 1 b) 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 1 a). To obtain the histogram, the 20 observations or samples, whose values are given in Table 1, are grouped into intervals 1 °C wide. (Preparation of a histogram is, of course, not required for the statistical analysis of the data.)

Table 1 — Twenty repeated observations of the temperature t grouped in 1 °C intervals

Interval $t_1 \leqslant t < t_2$		Temperature
t_1/°C	t_2/°C	t/°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	—

The arithmetic mean or average $\bar{t}$ of the $n = 20$ observations calculated according to Equation (3) is $\bar{t} = 100{,}145\ °C$ 100,14 °C and is assumed to be the best estimate of the expectation $\propto_t$ of t based on the available data. The experimental standard deviation $s(t_k)$ calculated from Equation (4) is $s(t_k) = 1{,}489\ °C$ 1,49 °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 $u(\bar{t}) = s(\bar{t}) = s(t_k)/\sqrt{20} = 0{,}333\ °C$ 0,33 °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.

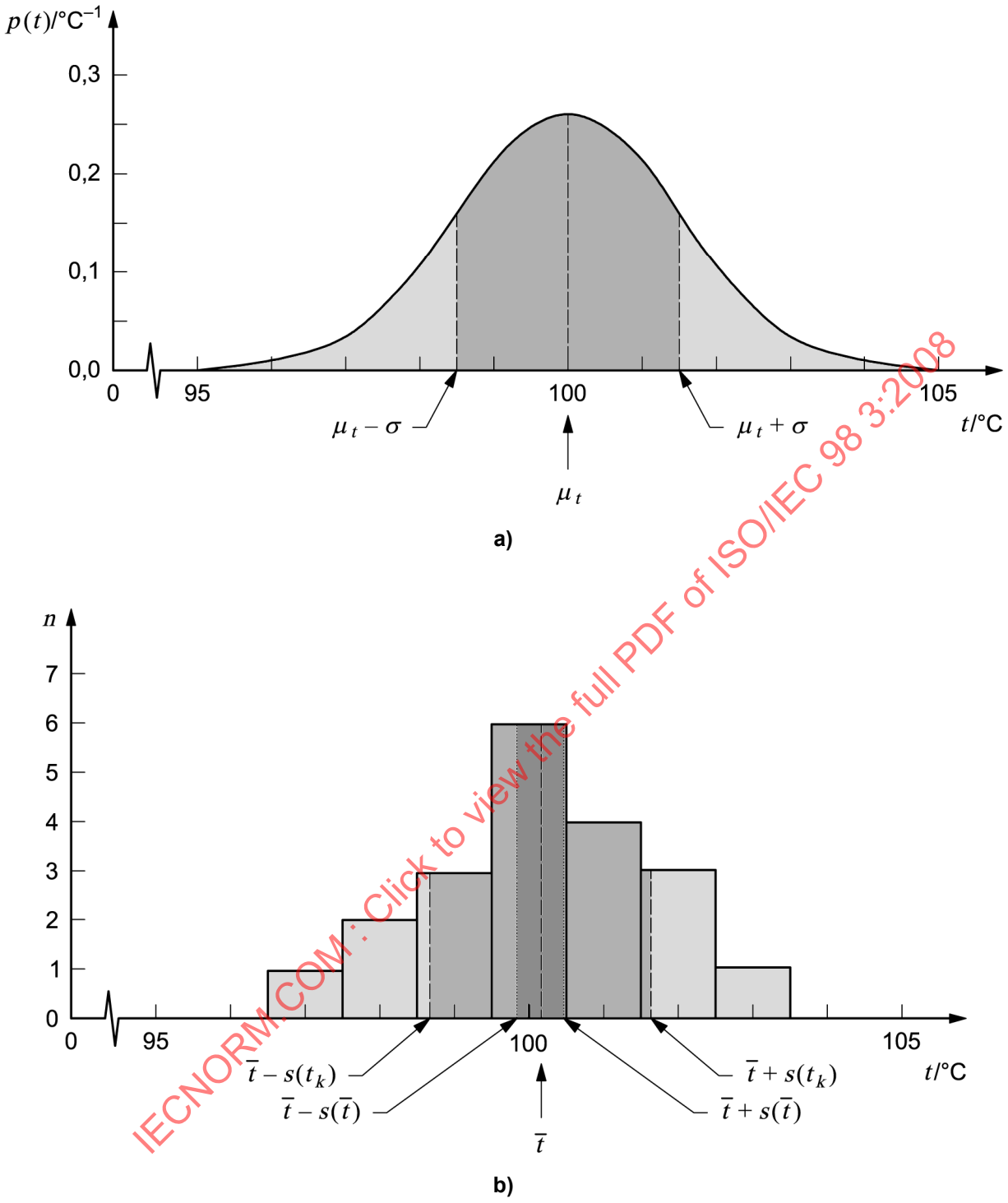


Figure 1 — Graphical illustration of evaluating the standard uncertainty of an input quantity from repeated observations

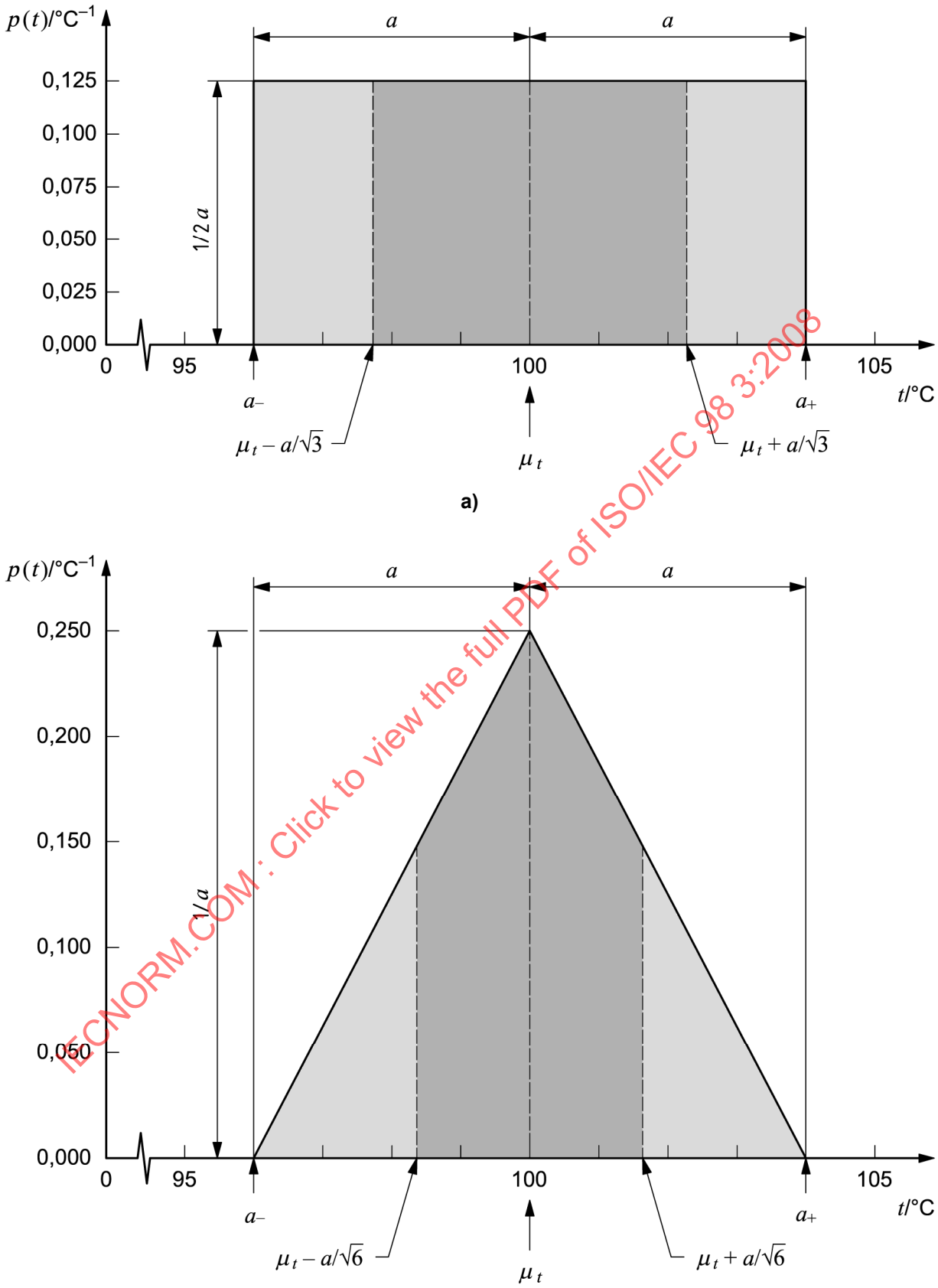


Figure 2 — Graphical illustration of evaluating the standard uncertainty of an input quantity from an *a priori* distribution

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 2 a), 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 $a_- = 96$ °C, upper bound $a_+ = 104$ °C, and thus half-width $a = (a_+ - a_-)/2 = 4$ °C (see 4.3.7). The probability density function of t is then

$$p(t) = 1/(2a), \qquad a_- \leqslant t \leqslant a_+$$

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

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

4.4.6 For the case illustrated in Figure 2 b), it is assumed that the available information concerning t is less limited and that t can be described by a symmetric, triangular *a priori* probability distribution of the same lower bound $a_- = 96$ °C, the same upper bound $a_+ = 104$ °C, and thus the same half-width $a = (a_+ - a_-)/2 = 4$ °C as in 4.4.5 (see 4.3.9). The probability density function of t is then

$$p(t) = (t - a_-)/a^2, \qquad a_- \leqslant t \leqslant (a_+ + a_-)/2$$

$$p(t) = (a_+ - t)/a^2, \qquad (a_+ + a_-)/2 \leqslant t \leqslant a_+$$

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

As indicated in 4.3.9, the expectation of t is $\propto_t = (a_+ + a_-)/2 = 100$ °C, which follows from C.3.1. The standard uncertainty of this estimate is $u(\propto_t) = a/\sqrt{6}$ 1,6 °C , which follows from C.3.2 [see Equation 9 b)].

The above value, $u(\propto_t) = 1{,}6$ °C, may be compared with $u(\propto_t) = 2{,}3$ °C obtained in 4.4.5 from a rectangular distribution of the same 8 °C width; with $\sigma = 1{,}5$ °C of the normal distribution of Figure 1 a) whose $-2{,}58\sigma$ to $+2{,}58\sigma$ width, which encompasses 99 percent of the distribution, is nearly 8 °C; and with $u(\bar{t}) = 0{,}33$ °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 $x_1, x_2, \ldots, x_N$ (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

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

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 $Y = f(X_1, X_2, ..., X_N)$, 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 normal, 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 $x_1, x_2, ..., x_N$. In particular, the change in y produced by a small change ∂x_i in input estimate x_i is given by $(\partial y)_i = (\partial f/\partial x_i)(\partial x_i)$. 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

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

where

$$c_i = \partial f / \partial x_i, \quad u_i(y) = |c_i| u(x_i) \tag{11b}$$

NOTE 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

$$\frac{\partial f}{\partial x_i} = \left. \frac{\partial f}{\partial X_i} \right|_{x_1, x_2, ..., x_N}$$

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

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

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,

$$c_1 = .P/.V = 2V/\left\{R_0\left[1+\alpha\left(t-t_0\right)\right]\right\} = 2P/V$$

$$c_2 = .P/.R_0 = -V^2/\left\{R_0^2\left[1+\alpha\left(t-t_0\right)\right]\right\} = -P/R_0$$

$$c_3 = .P/.\alpha = -V^2\left(t-t_0\right)/\left\{R_0\left[1+\alpha\left(t-t_0\right)\right]^2\right\} = -P\left(t-t_0\right)/\left[1+\alpha\left(t-t_0\right)\right]$$

$$c_4 = .P/.t = -V^2\alpha/\left\{R_0\left[1+\alpha\left(t-t_0\right)\right]^2\right\} = -P\alpha/\left[1+\alpha\left(t-t_0\right)\right]$$

and

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

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), $Y = Y_0 + c_1\delta_1 + c_2\delta_2 + \ldots + c_N\delta_N$, where $Y_0 = f(X_{1,0}, X_{2,0}, \ldots, X_{N,0})$, $c_i = (\partial f/\partial X_i)$ 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 δ_i (see E.3.1).

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

$$u_c^2(V) = u^2(\bar{V}) + u^2(.\bar{V}) = (12\ \mu\text{V})^2 + (8{,}7\ \mu\text{V})^2 = 219 \cdot 10^{-12}\ \text{V}^2$$

and the combined standard uncertainty is $u_c(V) = 15\ \mu\text{V}$, which corresponds to a relative combined standard uncertainty $u_c(V)/V$ of $16 \cdot 10^{-6}$ (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 $Y = c_1X_1 + c_2X_2 + \ldots + c_NX_N$ 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 $Y = cX_1^{p_1}X_2^{p_2}\ldots X_N^{p_N}$ and the exponents p_i are known positive or negative numbers having negligible uncertainties, the combined variance, Equation (10), can be expressed as

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

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 $|x_i| \neq 0$.]

NOTE 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: $Z = \ln c + \sum_{i=1}^{N} p_i W_i$.

NOTE 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 (11a) 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

$$u_c^2(y) = \sum_{i=1}^{N}\sum_{j=1}^{N}\frac{\partial f}{\partial x_i}\frac{\partial f}{\partial x_j}u(x_i, x_j) = \sum_{i=1}^{N}\left(\frac{\partial f}{\partial x_i}\right)^2 u^2(x_i) + 2\sum_{i=1}^{N-1}\sum_{j=i+1}^{N}\frac{\partial f}{\partial x_i}\frac{\partial f}{\partial x_j}u(x_i, x_j) \tag{13}$$

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)

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

where $r(x_i, x_j) = r(x_j, x_i)$, and $-1 \leqslant r(x_i, x_j) \leqslant +1$. 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

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

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

$$u_c^2(y) = \sum_{i=1}^{N}c_i^2u^2(x_i) + 2\sum_{i=1}^{N-1}\sum_{j=i+1}^{N}c_ic_ju(x_i)u(x_j)r(x_i, x_j) \tag{16}$$

NOTE 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

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

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 $R_i = 1\,000\ .\ ,$ are calibrated with a negligible uncertainty of comparison in terms of the same 1 000 . standard resistor R_s characterized by a standard uncertainty $u(R_s) = 100\ \text{m}.$ 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_{ref} of nominal value 10 k. . Thus $R_{ref} = f(R_i) = \sum_{i=1}^{10} R_i$. 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 $\partial f/\partial x_i = \partial R_{ref}/\partial R_i = 1$, 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_{ref}, $u_c(R_{ref}) = \sum_{i=1}^{10} u(R_s) = 10 \cdot (100\ \text{m}.\ ) = 1\ .\ .$ The result $u_c(R_{ref}) = \left[\sum_{i=1}^{10} u^2(R_s)\right]^{1/2} = 0{,}32\ .$ 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.

NOTE 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 $u_{ij} (i\ .\ j)$ 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.)

NOTE 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.

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

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

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 $\propto_q$ and $\propto_r$ 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

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

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): $r(x_i, x_j) = r(\bar{X}_i, \bar{X}_j) = s(\bar{X}_i, \bar{X}_j)/\left[s(\bar{X}_i)s(\bar{X}_j)\right]$.

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

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:

$$U = k u_c(y) \tag{18}$$

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 $y - U \leqslant Y \leqslant y + U$.

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, 6.3.3 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 $Y = y \pm U = y \pm ku_c(y)$ 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 $u_c(y)/|y|$, $|y|$. 0, when appropriate;

d) give the information outlined in 7.2.7 or refer to a published document that contains it.

If it is deemed useful for the intended users of the measurement result, for example, to aid in future calculations of coverage factors or to assist in understanding the measurement, one may indicate

. the estimated effective degrees of freedom ν_{eff} (see G.4);

. the Type A and Type B combined standard uncertainties $u_{cA}(y)$ and $u_{cB}(y)$ and their estimated effective degrees of freedom ν_{effA} and ν_{effB} (see G.4.1, Note 3).

7.2.2 When the measure of uncertainty is $u_c(y)$, it is preferable to state the numerical result of the measurement in one of the following four ways in order to prevent misunderstanding. (The quantity whose value is being reported is assumed to be a nominally 100 g standard of mass m_S; the words in parentheses may be omitted for brevity if u_c is defined elsewhere in the document reporting the result.)

1) “m_S = 100,021 47 g with (a combined standard uncertainty) u_c = 0,35 mg.”

2) “m_S = 100,021 47(35) g, where the number in parentheses is the numerical value of (the combined standard uncertainty) u_c referred to the corresponding last digits of the quoted result.”

3) "m_S = 100,021 47(0,000 35) g, where the number in parentheses is the numerical value of (the combined standard uncertainty) u_c expressed in the unit of the quoted result."

4) "m_S = (100,021 47 ± 0,000 35) g, where the number following the symbol ± is the numerical value of (the combined standard uncertainty) u_c and not a confidence interval."

NOTE The ± format should be avoided whenever possible because it has traditionally been used to indicate an interval corresponding to a high level of confidence and thus may be confused with expanded uncertainty (see 7.2.4). Further, although the purpose of the caveat in 4) is to prevent such confusion, writing $Y = y \pm u_c(y)$ might still be misunderstood to imply, especially if the caveat is accidentally omitted, that an expanded uncertainty with $k = 1$ is intended and that the interval $y - u_c(y) \leqslant Y \leqslant y + u_c(y)$ has a specified level of confidence p, namely, that associated with the normal distribution (see G.1.3). As indicated in 6.3.2 and Annex G, interpreting $u_c(y)$ in this way is usually difficult to justify.

7.2.3 When reporting the result of a measurement, and when the measure of uncertainty is the expanded uncertainty $U = ku_c(y)$, one should

a) give a full description of how the measurand Y is defined;

b) state the result of the measurement as $Y = y \pm U$ and give the units of y and U;

c) include the relative expanded uncertainty $U/|y|$, $|y| \neq 0$, when appropriate;

d) give the value of k used to obtain U [or, for the convenience of the user of the result, give both k and $u_c(y)$];

e) give the approximate level of confidence associated with the interval $y \pm U$ and state how it was determined;

f) give the information outlined in 7.2.7 or refer to a published document that contains it.

7.2.4 When the measure of uncertainty is U, it is preferable, for maximum clarity, to state the numerical result of the measurement as in the following example. (The words in parentheses may be omitted for brevity if U, u_c, and k are defined elsewhere in the document reporting the result.)

> "m_S = (100,021 47 ± 0,000 79) g, where the number following the symbol ± is the numerical value of (an expanded uncertainty) $U = ku_c$, with U determined from (a combined standard uncertainty) u_c = 0,35 mg and (a coverage factor) k = 2,26 based on the t-distribution for $\nu = 9$ degrees of freedom, and defines an interval estimated to have a level of confidence of 95 percent."

7.2.5 If a measurement determines simultaneously more than one measurand, that is, if it provides two or more output estimates y_i (see H.2, H.3, and H.4), then, in addition to giving y_i and $u_c(y_i)$, give the covariance matrix elements $u(y_i, y_j)$ or the elements $r(y_i, y_j)$ of the **correlation coefficient matrix** (C.3.6, Note 2) (and preferably both).

7.2.6 The numerical values of the estimate y and its standard uncertainty $u_c(y)$ or expanded uncertainty U should not be given with an excessive number of digits. It usually suffices to quote $u_c(y)$ and U [as well as the standard uncertainties $u(x_i)$ of the input estimates x_i] to at most two significant digits, although in some cases it may be necessary to retain additional digits to avoid round-off errors in subsequent calculations.

In reporting final results, it may sometimes be appropriate to round uncertainties up rather than to the nearest digit. For example, $u_c(y)$ = 10,47 m. might be rounded up to 11 m. . However, common sense should prevail and a value such as $u(x_i)$ = 28,05 kHz should be rounded down to 28 kHz. Output and input estimates should be rounded to be consistent with their uncertainties; for example, if y = 10,057 62 . with $u_c(y)$ = 27 m. , y should be rounded to 10,058 . . Correlation coefficients should be given with three-digit accuracy if their absolute values are near unity.

7.2.7 In the detailed report that describes how the result of a measurement and its uncertainty were obtained, one should follow the recommendations of 7.1.4 and thus

a) give the value of each input estimate x_i and its standard uncertainty $u(x_i)$ together with a description of how they were obtained;

b) give the estimated covariances or estimated correlation coefficients (preferably both) associated with all input estimates that are correlated, and the methods used to obtain them;

c) give the degrees of freedom for the standard uncertainty of each input estimate and how it was obtained;

d) give the functional relationship $Y = f(X_1, X_2, ..., X_N)$ and, when they are deemed useful, the partial derivatives or sensitivity coefficients $\partial f/\partial x_i$. However, any such coefficients determined experimentally should be given.

NOTE Since the functional relationship f may be extremely complex or may not exist explicitly but only as a computer program, it may not always be possible to give f and its derivatives. The function f may then be described in general terms or the program used may be cited by an appropriate reference. In such cases, it is important that it be clear how the estimate y of the measurand Y and its combined standard uncertainty $u_c(y)$ were obtained.

8 Summary of procedure for evaluating and expressing uncertainty

The steps to be followed for evaluating and expressing the uncertainty of the result of a measurement as presented in this *Guide* may be summarized as follows:

1) Express mathematically the relationship between the measurand Y and the input quantities X_i on which Y depends: $Y = f(X_1, X_2, ..., X_N)$. The function f should contain every quantity, including all corrections and correction factors, that can contribute a significant component of uncertainty to the result of the measurement (see 4.1.1 and 4.1.2).

2) Determine x_i, the estimated value of input quantity X_i, either on the basis of the statistical analysis of series of observations or by other means (see 4.1.3).

3) Evaluate the *standard uncertainty* $u(x_i)$ of each input estimate x_i. For an input estimate obtained from the statistical analysis of series of observations, the standard uncertainty is evaluated as described in 4.2 (*Type A evaluation of standard uncertainty*). For an input estimate obtained by other means, the standard uncertainty $u(x_i)$ is evaluated as described in 4.3 (*Type B evaluation of standard uncertainty*).

4) Evaluate the covariances associated with any input estimates that are correlated (see 5.2).

5) Calculate the result of the measurement, that is, the estimate y of the measurand Y, from the functional relationship f using for the input quantities X_i the estimates x_i obtained in step 2 (see 4.1.4).

6) Determine the *combined standard uncertainty* $u_c(y)$ of the measurement result y from the standard uncertainties and covariances associated with the input estimates, as described in Clause 5. If the measurement determines simultaneously more than one output quantity, calculate their covariances (see 7.2.5, H.2, H.3, and H.4).

7) If it is necessary to give an *expanded uncertainty* U, whose purpose is to provide an interval $y - U$ to $y + U$ that may be expected to encompass a large fraction of the distribution of values that could reasonably be attributed to the measurand Y, multiply the combined standard uncertainty $u_c(y)$ by a *coverage factor* k, typically in the range 2 to 3, to obtain $U = ku_c(y)$. Select k on the basis of the level of confidence required of the interval (see 6.2, 6.3, and especially Annex G, which discusses the selection of a value of k, that produces an interval having a level of confidence close to a specified value).

8) Report the result of the measurement y together with its combined standard uncertainty $u_c(y)$ or expanded uncertainty U as discussed in 7.2.1 and 7.2.3; use one of the formats recommended in 7.2.2 and 7.2.4. Describe, as outlined also in Clause 7, how y and $u_c(y)$ or U were obtained.

Annex A

Recommendations of Working Group and CIPM

A.1 Recommendation INC-1 (1980)

The Working Group on the Statement of Uncertainties (see Foreword) was convened in October 1980 by the Bureau International des Poids et Mesures (BIPM) in response to a request of the Comité International des Poids et Mesures (CIPM). It prepared a detailed report for consideration by the CIPM that concluded with Recommendation INC-1 (1980) [2]. The English translation of this Recommendation is given in 0.7 of this *Guide* and the French text, which is authoritative, is as follows [2]:

Expression des incertitudes expérimentales

Recommandation INC-1 (1980)

1) L'incertitude d'un résultat de mesure comprend généralement plusieurs composantes qui peuvent être groupées en deux catégories d'après la méthode utilisée pour estimer leur valeur numérique:

   A. celles qui sont évaluées à l'aide de méthodes statistiques,

   B. celles qui sont évaluées par d'autres moyens.

   Il n'y a pas toujours une correspondance simple entre le classement dans les catégories A ou B et le caractère «aléatoire» ou «systématique» utilisé antérieurement pour classer les incertitudes. L'expression «incertitude systématique» est susceptible de conduire à des erreurs d'interprétation; elle doit être évitée.

   Toute description détaillée de l'incertitude devrait comprendre une liste complète de ses composantes et indiquer pour chacune la méthode utilisée pour lui attribuer une valeur numérique.

2) Les composantes de la catégorie A sont caractérisées par les variances estimées s_i^2 (ou les «écart-types» estimés s_i) et les nombres v_i de degrés de liberté. Le cas échéant, les covariances estimées doivent être données.

3) Les composantes de la catégorie B devraient être caractérisées par des termes u_j^2 qui puissent être considérés comme des approximations des variances correspondantes dont on admet l'existence. Les termes u_j^2 peuvent être traités comme des variances et les termes u_j comme des écarts-types. Le cas échéant, les covariances doivent être traitées de façon analogue.

4) L'incertitude composée devrait être caractérisée par la valeur obtenue en appliquant la méthode usuelle de combinaison des variances. L'incertitude composée ainsi que ses composantes devraient être exprimées sous la forme d'«écart-types».

5) Si pour des utilisations particulières on est amené à multiplier par un facteur l'incertitude composée afin d'obtenir une incertitude globale, la valeur numérique de ce facteur doit toujours être donnée.

A.2 Recommendation 1 (CI-1981)

The CIPM reviewed the report submitted to it by the Working Group on the Statement of Uncertainties and adopted the following recommendation at its 70th meeting held in October 1981 [3]:

Recommendation 1 (CI-1981)

Expression of experimental uncertainties

The Comité International des Poids et Mesures

considering

. the need to find an agreed way of expressing measurement uncertainty in metrology,

. the effort that has been devoted to this by many organizations over many years,

. the encouraging progress made in finding an acceptable solution, which has resulted from the discussions of the Working Group on the Expression of Uncertainties which met at BIPM in 1980,

recognizes

. that the proposals of the Working Group might form the basis of an eventual agreement on the expression of uncertainties,

recommends

. that the proposals of the Working Group be diffused widely;

. that BIPM attempt to apply the principles therein to international comparisons carried out under its auspices in the years to come;

. that other interested organizations be encouraged to examine and test these proposals and let their comments be known to BIPM;

. that after two or three years BIPM report back on the application of these proposals.

A.3 Recommendation 1 (CI-1986)

The CIPM further considered the matter of the expression of uncertainties at its 75th meeting held in October 1986 and adopted the following recommendation [4]:

Recommendation 1 (CI-1986)

Expression of uncertainties in work carried out under the auspices of the CIPM

The Comité International des Poids et Mesures,

considering the adoption by the Working Group on the Statement of Uncertainties of Recommendation INC-1 (1980) and the adoption by the CIPM of Recommendation 1 (CI-1981),

considering that certain members of Comités Consultatifs may want clarification of this Recommendation for the purposes of work that falls under their purview, especially for international comparisons,

recognizes that paragraph 5 of Recommendation INC-1 (1980) relating to particular applications, especially those having commercial significance, is now being considered by a working group of the International Standards Organization (ISO) common to the ISO, OIML and IEC, with the concurrence and cooperation of the CIPM,

requests that paragraph 4 of Recommendation INC-1 (1980) should be applied by all participants in giving the results of all international comparisons or other work done under the auspices of the CIPM and the Comités Consultatifs and that the combined uncertainty of type A and type B uncertainties in terms of *one standard deviation* should be given.

Annex B

General metrological terms

B.1 Source of definitions

The definitions of the general metrological terms relevant to this *Guide* that are given here have been taken from the *International vocabulary of basic and general terms in metrology* (abbreviated VIM), second edition, 1993* [6], published by the International Organization for Standardization (ISO), in the name of the seven organizations that supported its development and nominated the experts who prepared it: the Bureau International des Poids et Mesures (BIPM), the International Electrotechnical Commission (IEC), the International Federation of Clinical Chemistry (IFCC), ISO, the International Union of Pure and Applied Chemistry (IUPAC), the International Union of Pure and Applied Physics (IUPAP), and the International Organization of Legal Metrology (OIML). The VIM should be the first source consulted for the definitions of terms not included either here or in the text.

NOTE Some basic statistical terms and concepts are given in Annex C, while the terms "true value", "error", and "uncertainty" are further discussed in Annex D.

B.2 Definitions

As in Clause 0, in the definitions that follow, the use of parentheses around certain words of some terms means that the words may be omitted if this is unlikely to cause confusion.

The terms in boldface in some notes are additional metrological terms defined in those notes, either explicitly or implicitly (see Reference [6]).

B.2.1
(measurable) quantity
attribute of a phenomenon, body or substance that may be distinguished qualitatively and determined quantitatively

NOTE 1 The term quantity may refer to a quantity in a general sense (see Example 1) or to a **particular quantity** (see Example 2).

EXAMPLE 1 Quantities in a general sense: length, time, mass, temperature, electrical resistance, amount-of-substance concentration.

EXAMPLE 2 Particular quantities:

— length of a given rod

— electrical resistance of a given specimen of wire

— amount-of-substance concentration of ethanol in a given sample of wine.

NOTE 2 Quantities that can be placed in order of magnitude relative to one another are called **quantities of the same kind**.

* **Footnote to the 2008 version:**
The third edition of the vocabulary was published in 2007, under the title ISO/IEC Guide 99, *International vocabulary of metrology — Basic and general concepts and associated terms (VIM)*.

NOTE 3 Quantities of the same kind may be grouped together into **categories of quantities**, for example:

— work, heat, energy

— thickness, circumference, wavelength.

NOTE 4 **Symbols for quantities** are given in ISO 31*.

[VIM:1993, definition 1.1]

B.2.2
value (of a quantity)
magnitude of a particular quantity generally expressed as a unit of measurement multiplied by a number

EXAMPLE 1 Length of a rod: 5,34 m or 534 cm.

EXAMPLE 2 Mass of a body: 0,152 kg or 152 g.

EXAMPLE 3 Amount of substance of a sample of water (H_2O): 0,012 mol or 12 mmol.

NOTE 1 The value of a quantity may be positive, negative or zero.

NOTE 2 The value of a quantity may be expressed in more than one way.

NOTE 3 The values of quantities of dimension one are generally expressed as pure numbers.

NOTE 4 A quantity that cannot be expressed as a unit of measurement multiplied by a number may be expressed by reference to a conventional reference scale or to a measurement procedure or to both.

[VIM:1993, definition 1.18]

B.2.3
true value (of a quantity)
value consistent with the definition of a given particular quantity

NOTE 1 This is a value that would be obtained by a perfect measurement.

NOTE 2 True values are by nature indeterminate.

NOTE 3 The indefinite article “a”, rather than the definite article “the”, is used in conjunction with “true value” because there may be many values consistent with the definition of a given particular quantity.

[VIM:1993, definition 1.19]

Guide Comment: See Annex D, in particular D.3.5, for the reasons why the term “true value” is not used in this *Guide* and why the terms “true value of a measurand” (or of a quantity) and “value of a measurand” (or of a quantity) are viewed as equivalent.

B.2.4
conventional true value (of a quantity)
value attributed to a particular quantity and accepted, sometimes by convention, as having an uncertainty appropriate for a given purpose

EXAMPLE 1 At a given location, the value assigned to the quantity realized by a reference standard may be taken as a conventional true value.

EXAMPLE 2 The CODATA (1986) recommended value for the Avogadro constant: $6{,}022\ 136\ 7 \cdot 10^{23}\ \text{mol}^{-1}$.

---

* **Footnote to the 2008 version:**
The ISO 31 series is under revision as a series of ISO 80000 and IEC 80000 documents. (Some of these documents have already been published.)

NOTE 1 "Conventional true value" is sometimes called **assigned value**, **best estimate** of the value, **conventional value** or **reference value**. "Reference value", in this sense, should not be confused with "reference value" in the sense used in the Note to VIM:1993, definition 5.7.

NOTE 2 Frequently, a number of results of measurements of a quantity is used to establish a conventional true value.

[VIM:1993, definition 1.20]

Guide Comment: See the *Guide* Comment to B.2.3.

B.2.5
measurement
set of operations having the object of determining a value of a quantity

NOTE The operations may be performed automatically.

[VIM:1993, definition 2.1]

B.2.6
principle of measurement
scientific basis of a measurement

EXAMPLE 1 The thermoelectric effect applied to the measurement of temperature.

EXAMPLE 2 The Josephson effect applied to the measurement of electric potential difference.

EXAMPLE 3 The Doppler effect applied to the measurement of velocity.

EXAMPLE 4 The Raman effect applied to the measurement of the wave number of molecular vibrations.

[VIM:1993, definition 2.3]

B.2.7
method of measurement
logical sequence of operations, described generically, used in the performance of measurements

NOTE Methods of measurement may be qualified in various ways such as:

— substitution method
— differential method
— null method.

[VIM:1993, definition 2.4]

B.2.8
measurement procedure
set of operations, described specifically, used in the performance of particular measurements according to a given method

NOTE A measurement procedure is usually recorded in a document that is sometimes itself called a "measurement procedure" (or a **measurement method**) and is usually in sufficient detail to enable an operator to carry out a measurement without additional information.

[VIM:1993, definition 2.5]

B.2.9
measurand
particular quantity subject to measurement

EXAMPLE Vapour pressure of a given sample of water at 20 °C.

NOTE The specification of a measurand may require statements about quantities such as time, temperature and pressure.

[VIM:1993, definition 2.6]

B.2.10
influence quantity
quantity that is not the measurand but that affects the result of the measurement

EXAMPLE 1 Temperature of a micrometer used to measure length.

EXAMPLE 2 Frequency in the measurement of the amplitude of an alternating electric potential difference.

EXAMPLE 3 Bilirubin concentration in the measurement of haemoglobin concentration in a sample of human blood plasma.

[VIM:1993, definition 2.7]

Guide Comment: The definition of influence quantity is understood to include values associated with measurement standards, reference materials, and reference data upon which the result of a measurement may depend, as well as phenomena such as short-term measuring instrument fluctuations and quantities such as ambient temperature, barometric pressure and humidity.

B.2.11
result of a measurement
value attributed to a measurand, obtained by measurement

NOTE 1 When a result is given, it should be made clear whether it refers to:

— the indication

— the uncorrected result

— the corrected result

and whether several values are averaged.

NOTE 2 A complete statement of the result of a measurement includes information about the uncertainty of measurement.

[VIM:1993, definition 3.1]

B.2.12
uncorrected result
result of a measurement before correction for systematic error

[VIM:1993, definition 3.3]

B.2.13
corrected result
result of a measurement after correction for systematic error

[VIM:1993, definition 3.4]

B.2.14
accuracy of measurement
closeness of the agreement between the result of a measurement and a true value of the measurand

NOTE 1 "Accuracy" is a qualitative concept.

NOTE 2 The term **precision** should not be used for "accuracy".

[VIM:1993, definition 3.5]

Guide Comment: See the *Guide* Comment to B.2.3.

B.2.15
repeatability (of results of measurements)
closeness of the agreement between the results of successive measurements of the same measurand carried out under the same conditions of measurement

NOTE 1 These conditions are called **repeatability conditions**.

NOTE 2 Repeatability conditions include:

— the same measurement procedure
— the same observer
— the same measuring instrument, used under the same conditions
— the same location
— repetition over a short period of time.

NOTE 3 Repeatability may be expressed quantitatively in terms of the dispersion characteristics of the results.

[VIM:1993, definition 3.6]

B.2.16
reproducibility (of results of measurements)
closeness of the agreement between the results of measurements of the same measurand carried out under changed conditions of measurement

NOTE 1 A valid statement of reproducibility requires specification of the conditions changed.

NOTE 2 The changed conditions may include:

— principle of measurement
— method of measurement
— observer
— measuring instrument
— reference standard
— location
— conditions of use
— time.

NOTE 3 Reproducibility may be expressed quantitatively in terms of the dispersion characteristics of the results.

NOTE 4 Results are here usually understood to be corrected results.

[VIM:1993, definition 3.7]

B.2.17
experimental standard deviation
for a series of n measurements of the same measurand, the quantity $s(q_k)$ characterizing the dispersion of the results and given by the formula:

$$s(q_k) = \sqrt{\frac{\sum_{j=1}^{n}(q_j - \bar{q})^2}{n-1}}$$

q_k being the result of the kth measurement and $\bar{q}$ being the arithmetic mean of the n results considered

NOTE 1 Considering the series of n values as a sample of a distribution, $\bar{q}$ is an unbiased estimate of the mean μ_q, and $s^2(q_k)$ is an unbiased estimate of the variance σ^2, of that distribution.

NOTE 2 The expression $s(q_k)/\sqrt{n}$ is an estimate of the standard deviation of the distribution of $\bar{q}$ and is called the **experimental standard deviation of the mean**.

NOTE 3 "Experimental standard deviation of the mean" is sometimes incorrectly called **standard error of the mean**.

NOTE 4 Adapted from VIM:1993, definition 3.8.

Guide Comment: Some of the symbols used in the VIM have been changed in order to achieve consistency with the notation used in 4.2 of this *Guide*.

B.2.18
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

NOTE 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.

NOTE 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 can also be characterized by standard deviations, are evaluated from assumed probability distributions based on experience or other information.

NOTE 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.

[VIM:1993, definition 3.9]

Guide Comment: It is pointed out in the VIM that this definition and the notes are identical to those in this *Guide* (see 2.2.3).

B.2.19
error (of measurement)
result of a measurement minus a true value of the measurand

NOTE 1 Since a true value cannot be determined, in practice a conventional true value is used [see VIM:1993, definitions 1.19 (B.2.3) and 1.20 (B.2.4)].

NOTE 2 When it is necessary to distinguish "error" from "relative error", the former is sometimes called **absolute error of measurement**. This should not be confused with **absolute value of error**, which is the modulus of the error.

[VIM:1993, definition 3.10]

Guide Comment: If the result of a measurement depends on the values of quantities other than the measurand, the errors of the measured values of these quantities contribute to the error of the result of the measurement. Also see the *Guide* Comment to B.2.22 and to B.2.3.

B.2.20
relative error
error of measurement divided by a true value of the measurand

NOTE Since a true value cannot be determined, in practice a conventional true value is used [see VIM:1993, definitions 1.19 (B.2.3) and 1.20 (B.2.4)].

[VIM:1993, definition 3.12]

Guide Comment: See the *Guide* Comment to B.2.3.

B.2.21
random error
result of a measurement minus the mean that would result from an infinite number of measurements of the same measurand carried out under repeatability conditions

NOTE 1 Random error is equal to error minus systematic error.

NOTE 2 Because only a finite number of measurements can be made, it is possible to determine only an estimate of random error.

[VIM:1993, definition 3.13]

Guide Comment: See the *Guide* Comment to B.2.22.

B.2.22
systematic error
mean that would result from an infinite number of measurements of the same measurand carried out under repeatability conditions minus a true value of the measurand

NOTE 1 Systematic error is equal to error minus random error.

NOTE 2 Like true value, systematic error and its causes cannot be completely known.

NOTE 3 For a measuring instrument, see “bias” (VIM:1993, definition 5.25).

[VIM:1993, definition 3.14]

Guide Comment: The error of the result of a measurement (see B.2.19) may often be considered as arising from a number of random and systematic effects that contribute individual components of error to the error of the result. Also see the *Guide* Comment to B.2.19 and to B.2.3.

B.2.23
correction
value added algebraically to the uncorrected result of a measurement to compensate for systematic error

NOTE 1 The correction is equal to the negative of the estimated systematic error.

NOTE 2 Since the systematic error cannot be known perfectly. the compensation cannot be complete.

[VIM:1993, definition 3.15]

B.2.24
correction factor
numerical factor by which the uncorrected result of a measurement is multiplied to compensate for systematic error

NOTE Since the systematic error cannot be known perfectly, the compensation cannot be complete.

[VIM:1993, definition 3.16]

Annex C

Basic statistical terms and concepts

C.1 Source of definitions

The definitions of the basic statistical terms given in this annex are taken from International Standard ISO 3534-1:1993* [7]. This should be the first source consulted for the definitions of terms not included here. Some of these terms and their underlying concepts are elaborated upon in C.3 following the presentation of their formal definitions in C.2 in order to facilitate further the use of this *Guide*. However, C.3, which also includes the definitions of some related terms, is not based directly on ISO 3534-1:1993.

C.2 Definitions

As in Clause 0 and Annex B, the use of parentheses around certain words of some terms means that the words may be omitted if this is unlikely to cause confusion.

Terms C.2.1 to C.2.14 are defined in terms of the properties of populations. The definitions of terms C.2.15 to C.2.31 are related to a set of observations (see Reference [7]).

C.2.1
probability
a real number in the scale 0 to 1 attached to a random event

NOTE It can be related to a long-run relative frequency of occurrence or to a degree of belief that an event will occur. For a high degree of belief, the probability is near 1.

[ISO 3534-1:1993, definition 1.1]

C.2.2
random variable
variate
a variable that may take any of the values of a specified set of values and with which is associated a *probability distribution* [ISO 3534-1:1993, definition 1.3 (C.2.3)]

NOTE 1 A random variable that may take only isolated values is said to be “discrete”. A random variable which may take any value within a finite or infinite interval is said to be “continuous”.

NOTE 2 The probability of an event A is denoted by Pr(A) or P(A).

[ISO 3534-1:1993, definition 1.2]

Guide Comment: The symbol Pr(A) is used in this *Guide* in place of the symbol P_r(A) used in ISO 3534-1:1993.

* **Footnote to the 2008 version:**
ISO 3534-1:1993 has been cancelled and replaced by ISO 3534-1:2006. Note that some of the terms and definitions have been revised. For further information, see the latest edition.

C.2.3
probability distribution (of a random variable)
a function giving the probability that a random variable takes any given value or belongs to a given set of values

NOTE The probability on the whole set of values of the random variable equals 1.

[ISO 3534-1:1993, definition 1.3]

C.2.4
distribution function
a function giving, for every value x, the probability that the random variable X be less than or equal to x:

$$F(x) = \Pr(X \leqslant x)$$

[ISO 3534-1:1993, definition 1.4]

C.2.5
probability density function (for a continuous random variable)
the derivative (when it exists) of the distribution function:

$$f(x) = \mathrm{d}F(x)/\mathrm{d}x$$

NOTE $f(x)\,\mathrm{d}x$ is the "probability element":

$$f(x)\,\mathrm{d}x = \Pr(x < X < x + \mathrm{d}x)$$

[ISO 3534-1:1993, definition 1.5]

C.2.6
probability mass function
a function giving, for each value x_i of a discrete random variable X, the probability p_i that the random variable equals x_i:

$$p_i = \Pr(X = x_i)$$

[ISO 3534-1:1993, definition 1.6]

C.2.7
parameter
a quantity used in describing the probability distribution of a random variable

[ISO 3534-1:1993, definition 1.12]

C.2.8
correlation
the relationship between two or several random variables within a distribution of two or more random variables

NOTE Most statistical measures of correlation measure only the degree of linear relationship.

[ISO 3534-1:1993, definition 1.13]

C.2.9
expectation (of a random variable or of a probability distribution)
expected value
mean

1) For a discrete random variable X taking the values x_i with the probabilities p_i, the expectation, if it exists, is

$$\propto = E(X) = \sum p_i x_i$$

the sum being extended over all the values x_i which can be taken by X.

2) For a continuous random variable X having the probability density function $f(x)$, the expectation, if it exists, is

$$\propto = E(X) = \int x f(x)\,\mathrm{d}x$$

the integral being extended over the interval(s) of variation of X.

[ISO 3534-1:1993, definition 1.18]

C.2.10
centred random variable
a random variable the expectation of which equals zero

NOTE If the random variable X has an expectation equal to $\propto$, the corresponding centred random variable is $(X - \propto)$.

[ISO 3534-1:1993, definition 1.21]

C.2.11
variance (of a random variable or of a probability distribution)
the expectation of the square of the *centred random variable* [ISO 3534-1:1993, definition 1.21 (C.2.10)]:

$$\sigma^2 = V(X) = E\left\{\left[X - E(X)\right]^2\right\}$$

[ISO 3534-1:1993, definition 1.22]

C.2.12
standard deviation (of a random variable or of a probability distribution)
the positive square root of the variance:

$$\sigma = \sqrt{V(X)}$$

[ISO 3534-1:1993, definition 1.23]

C.2.13
central moment [2] **of order** q
in a univariate distribution, the expectation of the qth power of the centred random variable $(X - \propto)$:

$$E\left[(X - \propto)^q\right]$$

NOTE The central moment of order 2 is the *variance* [ISO 3534-1:1993, definition 1.22 (C.2.11)] of the random variable X.

[ISO 3534-1:1993, definition 1.28]

2) If, in the definition of the moments, the quantities X, $X - a$, Y, $Y - b$, etc. are replaced by their absolute values, i.e. $|X|$, $|X - a|$, $|Y|$, $|Y - b|$, etc., other moments called "absolute moments" are defined.

C.2.14
normal distribution
Laplace-Gauss distribution
the probability distribution of a continuous random variable X, the probability density function of which is

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left[-\frac{1}{2}\left(\frac{x-\propto}{\sigma}\right)^2\right]$$

for $-8 < x < +8$.

NOTE $\propto$ is the expectation and σ is the standard deviation of the normal distribution.

[ISO 3534-1:1993, definition 1.37]

C.2.15
characteristic
a property which helps to identify or differentiate between items of a given population

NOTE The characteristic may be either quantitative (by variables) or qualitative (by attributes).

[ISO 3534-1:1993, definition 2.2]

C.2.16
population
the totality of items under consideration

NOTE In the case of a random variable, the *probability distribution* [ISO 3534-1:1993, definition 1.3 (C.2.3)] is considered to define the population of that variable.

[ISO 3534-1:1993, definition 2.3]

C.2.17
frequency
the number of occurrences of a given type of event or the number of observations falling into a specified class

[ISO 3534-1:1993, definition 2.11]

C.2.18
frequency distribution
the empirical relationship between the values of a characteristic and their frequencies or their relative frequencies

NOTE The distribution may be graphically presented as a *histogram* (ISO 3534-1:1993, definition 2.17), *bar chart* (ISO 3534-1:1993, definition 2.18), *cumulative frequency polygon* (ISO 3534-1:1993, definition 2.19), or as a *two-way table* (ISO 3534-1:1993, definition 2.22).

[ISO 3534-1:1993, definition 2.15]

C.2.19
arithmetic mean
average
the sum of values divided by the number of values

NOTE 1 The term “mean” is used generally when referring to a population parameter and the term “average” when referring to the result of a calculation on the data obtained in a sample.

NOTE 2 The average of a simple random sample taken from a population is an unbiased estimator of the mean of this population. However, other estimators, such as the geometric or harmonic mean, or the median or mode, are sometimes used.

[ISO 3534-1:1993, definition 2.26]

C.2.20
variance
a measure of dispersion, which is the sum of the squared deviations of observations from their average divided by one less than the number of observations

EXAMPLE For n observations x_1, x_2, ..., x_n with average

$$\bar{x} = (1/n)\sum x_i$$

the variance is

$$s^2 = \frac{1}{n-1}\sum (x_i - \bar{x})^2$$

NOTE 1 The sample variance is an unbiased estimator of the population variance.

NOTE 2 The variance is $n/(n-1)$ times the central moment of order 2 (see note to ISO 3534-1:1993, definition 2.39).

[ISO 3534-1:1993, definition 2.33]

Guide Comment: The variance defined here is more appropriately designated the "sample estimate of the population variance". The variance of a sample is usually defined to be the central moment of order 2 of the sample (see C.2.13 and C.2.22).

C.2.21
standard deviation
the positive square root of the variance

NOTE The sample standard deviation is a biased estimator of the population standard deviation.

[ISO 3534-1:1993, definition 2.34]

C.2.22
central moment of order q
in a distribution of a single characteristic, the arithmetic mean of the qth power of the difference between the observed values and their average $\bar{x}$:

$$\frac{1}{n}\sum_i (x_i - \bar{x})^q$$

where n is the number of observations

NOTE The central moment of order 1 is equal to zero.

[ISO 3534-1:1993, definition 2.37]

C.2.23
statistic
a function of the sample random variables

NOTE A statistic, as a function of random variables, is also a random variable and as such it assumes different values from sample to sample. The value of the statistic obtained by using the observed values in this function may be used in a statistical test or as an estimate of a population parameter, such as a mean or a standard deviation.

[ISO 3534-1:1993, definition 2.45]

C.2.24
estimation
the operation of assigning, from the observations in a sample, numerical values to the parameters of a distribution chosen as the statistical model of the population from which this sample is taken

NOTE A result of this operation may be expressed as a single value [point estimate; see ISO 3534-1:1993, definition 2.51 (C.2.26)] or as an interval estimate [see ISO 3534-1:1993, definitions 2.57 (C.2.27) and 2.58 (C.2.28)].

[ISO 3534-1:1993, definition 2.49]

C.2.25
estimator
a statistic used to estimate a population parameter

[ISO 3534-1:1993, definition 2.50]

C.2.26
estimate
the value of an estimator obtained as a result of an estimation

[ISO 3534-1:1993, definition 2.51]

C.2.27
two-sided confidence interval
when T_1 and T_2 are two functions of the observed values such that, θ being a population parameter to be estimated, the probability $\Pr(T_1 \leqslant . \leqslant T_2)$ is at least equal to $(1 - \alpha)$ [where $(1 - \alpha)$ is a fixed number, positive and less than 1], the interval between T_1 and T_2 is a two-sided $(1 - \alpha)$ confidence interval for $.$

NOTE 1 The limits T_1 and T_2 of the confidence interval are *statistics* [ISO 3534-1:1993, definition 2.45 (C.2.23)] and as such will generally assume different values from sample to sample.

NOTE 2 In a long series of samples, the relative frequency of cases where the true value of the population parameter $.$ is covered by the confidence interval is greater than or equal to $(1 - \alpha)$.

[ISO 3534-1:1993, definition 2.57]

C.2.28
one-sided confidence interval
when T is a function of the observed values such that, $.$ being a population parameter to be estimated, the probability $\Pr(T \geqslant .)$ [or the probability $\Pr(T \leqslant .)$] is at least equal to $(1 - \alpha)$ [where $(1 - \alpha)$ is a fixed number, positive and less than 1], the interval from the smallest possible value of $.$ up to T (or the interval from T up to the largest possible value of $.$) is a one-sided $(1 - \alpha)$ confidence interval for $.$

NOTE 1 The limit T of the confidence interval is a *statistic* [ISO 3534-1:1993, definition 2.45 (C.2.23)] and as such will generally assume different values from sample to sample.

NOTE 2 See Note 2 of ISO 3534-1:1993, definition 2.57 (C.2.27).

[ISO 3534-1:1993, definition 2.58]

C.2.29
confidence coefficient
confidence level
the value $(1 - \alpha)$ of the probability associated with a confidence interval or a statistical coverage interval [See ISO 3534-1:1993, definitions 2.57 (C.2.27), 2.58 (C.2.28) and 2.61 (C.2.30).]

NOTE $(1 - \alpha)$ is often expressed as a percentage.

[ISO 3534-1:1993, definition 2.59]

C.2.30
statistical coverage interval
an interval for which it can be stated with a given level of confidence that it contains at least a specified proportion of the population

NOTE 1 When both limits are defined by statistics, the interval is two-sided. When one of the two limits is not finite or consists of the boundary of the variable, the interval is one-sided.

NOTE 2 Also called “statistical tolerance interval”. This term should not be used because it may cause confusion with “tolerance interval” which is defined in ISO 3534-2:1993.

[ISO 3534-1:1993, definition 2.61]

C.2.31
degrees of freedom
in general, the number of terms in a sum minus the number of constraints on the terms of the sum

[ISO 3534-1:1993, definition 2.85]

C.3 Elaboration of terms and concepts

C.3.1 Expectation

The expectation of a function $g(z)$ over a probability density function $p(z)$ of the random variable z is defined by

$$E\left[g(z)\right] = \int g(z)\, p(z)\, \mathrm{d}z$$

where, from the definition of $p(z)$, $\int p(z)\,\mathrm{d}z = 1$. The expectation of the random variable z, denoted by $\propto_z$, and which is also termed the expected value or the mean of z, is given by

$$\propto_z = E(z) = \int z\, p(z)\, \mathrm{d}z$$

It is estimated statistically by $\bar{z}$, the arithmetic mean or average of n independent observations z_i of the random variable z, the probability density function of which is $p(z)$:

$$\bar{z} = \frac{1}{n}\sum_{i=1}^{n} z_i$$

C.3.2 Variance

The variance of a random variable is the expectation of its quadratic deviation about its expectation. Thus the variance of random variable z with probability density function $p(z)$ is given by

$$\sigma^2(z) = \int \left(z - \propto_z\right)^2 p(z)\, \mathrm{d}z$$

where $\propto_z$ is the expectation of z. The variance $\sigma^2(z)$ may be estimated by

$$s^2(z_i) = \frac{1}{n-1}\sum_{j=1}^{n}\left(z_j - \bar{z}\right)^2$$

where

$$\bar{z} = \frac{1}{n}\sum_{i=1}^{n} z_i$$

and the z_i are n independent observations of z.

NOTE 1 The factor $n - 1$ in the expression for $s^2(z_i)$ arises from the correlation between z_i and $\bar{z}$ and reflects the fact that there are only $n - 1$ independent items in the set $\{z_i - \bar{z}\}$.

NOTE 2 If the expectation $\propto_z$ of z is known, the variance may be estimated by

$$s^2(z_i) = \frac{1}{n}\sum_{i=1}^{n}(z_i - \propto_z)^2$$

The variance of the arithmetic mean or average of the observations, rather than the variance of the individual observations, is the proper measure of the uncertainty of a measurement result. The variance of a variable z should be carefully distinguished from the variance of the mean $\bar{z}$. The variance of the arithmetic mean of a series of n independent observations z_i of z is given by $\sigma^2(\bar{z}) = \sigma^2(z_i)/n$ and is estimated by the experimental variance of the mean

$$s^2(\bar{z}) = \frac{s^2(z_i)}{n} = \frac{1}{n(n-1)}\sum_{i=1}^{n}(z_i - \bar{z})^2$$

C.3.3 Standard deviation

The standard deviation is the positive square root of the variance. Whereas a Type A standard uncertainty is obtained by taking the square root of the statistically evaluated variance, it is often more convenient when determining a Type B standard uncertainty to evaluate a nonstatistical equivalent standard deviation first and then to obtain the equivalent variance by squaring the standard deviation.

C.3.4 Covariance

The covariance of two random variables is a measure of their mutual dependence. The covariance of random variables y and z is defined by

$$\mathrm{cov}(y, z) = \mathrm{cov}(z, y) = E\{[y - E(y)][z - E(z)]\}$$

which leads to

$$\begin{aligned}\mathrm{cov}(y, z) &= \mathrm{cov}(z, y)\\ &= \iint (y - \propto_y)(z - \propto_z)\, p(y, z)\, \mathrm{d}y\, \mathrm{d}z\\ &= \iint y\, z\, p(y, z)\, \mathrm{d}y\, \mathrm{d}z - \propto_y \propto_z\end{aligned}$$

where $p(y, z)$ is the joint probability density function of the two variables y and z. The covariance cov(y, z) [also denoted by $s(y, z)$] may be estimated by $s(y_i, z_i)$ obtained from n independent pairs of simultaneous observations y_i and z_i of y and z,

$$s(y_i, z_i) = \frac{1}{n-1}\sum_{j=1}^{n}(y_j - \bar{y})(z_j - \bar{z})$$

where

$$\bar{y} = \frac{1}{n}\sum_{i=1}^{n} y_i$$

and

$$\bar{z} = \frac{1}{n}\sum_{i=1}^{n} z_i$$

NOTE The estimated covariance of the two means $\bar{y}$ and $\bar{z}$ is given by $s(\bar{y}, \bar{z}) = s(y_i, z_i)/n$.

C.3.5 Covariance matrix

For a multivariate probability distribution, the matrix V with elements equal to the variances and covariances of the variables is termed the covariance matrix. The diagonal elements, $.(z, z) = \sigma^2(z)$ or $s(z_i, z_i) = s^2(z_i)$, are the variances, while the off-diagonal elements, $.(y, z)$ or $s(y_i, z_i)$, are the covariances.

C.3.6 Correlation coefficient

The correlation coefficient is a measure of the relative mutual dependence of two variables, equal to the ratio of their covariances to the positive square root of the product of their variances. Thus

$$.(y, z) = .(z, y) = \frac{.(y, z)}{\sqrt{.(y, y)\ .(z, z)}} = \frac{.(y, z)}{\sigma(y)\sigma(z)}$$

with estimates

$$r(y_i, z_i) = r(z_i, y_i) = \frac{s(y_i, z_i)}{\sqrt{s(y_i, y_i)s(z_i, z_i)}} = \frac{s(y_i, z_i)}{s(y_i)s(z_i)}$$

The correlation coefficient is a pure number such that $-1 \leqslant . \leqslant +1$ or $-1 \leqslant r(y_i, z_i) \leqslant +1$.

NOTE 1 Because $.$ and r are pure numbers in the range –1 to +1 inclusive, while covariances are usually quantities with inconvenient physical dimensions and magnitudes, correlation coefficients are generally more useful than covariances.

NOTE 2 For multivariate probability distributions, the correlation coefficient matrix is usually given in place of the covariance matrix. Since $.(y, y) = 1$ and $r(y_i, y_i) = 1$, the diagonal elements of this matrix are unity.

NOTE 3 If the input estimates x_i and x_j are correlated (see 5.2.2) and if a change δ_i in x_i produces a change δ_j in x_j, then the correlation coefficient associated with x_i and x_j is estimated approximately by

$$r(x_i, x_j) \quad u(x_i)\delta_j \Big/ \left[u(x_j)\delta_i\right]$$

This relation can serve as a basis for estimating correlation coefficients experimentally. It can also be used to calculate the approximate change in one input estimate due to a change in another if their correlation coefficient is known.

C.3.7 Independence

Two random variables are statistically independent if their joint probability distribution is the product of their individual probability distributions.

NOTE If two random variables are independent, their covariance and correlation coefficient are zero, but the converse is not necessarily true.

C.3.8 The t-distribution; Student's distribution

The t-distribution or Student's distribution is the probability distribution of a continuous random variable t whose probability density function is

$$p(t, \nu) = \frac{1}{\sqrt{\pi\nu}} \frac{\Gamma\left(\frac{\nu+1}{2}\right)}{\Gamma\left(\frac{\nu}{2}\right)} \left(1 + \frac{t^2}{\nu}\right)^{-(\nu+1)/2}, \qquad -8 < t < +8$$

where Γ is the gamma function and $\nu > 0$. The expectation of the t-distribution is zero and its variance is $\nu/(\nu - 2)$ for $\nu > 2$. As $\nu \to 8$, the t-distribution approaches a normal distribution with $\propto = 0$ and $\sigma = 1$ (see C.2.14).

The probability distribution of the variable $(\bar{z} - \propto_z)/s(\bar{z})$ is the t-distribution if the random variable z is normally distributed with expectation $\propto_z$, where $\bar{z}$ is the arithmetic mean of n independent observations z_i of z, $s(z_i)$ is the experimental standard deviation of the n observations, and $s(\bar{z}) = s(z_i)/\sqrt{n}$ is the experimental standard deviation of the mean $\bar{z}$ with $v = n - 1$ degrees of freedom.

Annex D

"True" value, error, and uncertainty

The term **true value** (B.2.3) has traditionally been used in publications on uncertainty but not in this *Guide* for the reasons presented in this annex. Because the terms "measurand", "error", and "uncertainty" are frequently misunderstood, this annex also provides additional discussion of the ideas underlying them to supplement the discussion given in Clause 3. Two figures are presented to illustrate why the concept of uncertainty adopted in this *Guide* is based on the measurement result and its evaluated uncertainty rather than on the unknowable quantities "true" value and error.

D.1 The measurand

D.1.1 The first step in making a measurement is to specify the measurand — the quantity to be measured; the measurand cannot be specified by a value but only by a description of a quantity. However, in principle, a measurand cannot be *completely* described without an infinite amount of information. Thus, to the extent that it leaves room for interpretation, incomplete definition of the measurand introduces into the uncertainty of the result of a measurement a component of uncertainty that may or may not be significant relative to the accuracy required of the measurement.

D.1.2 Commonly, the definition of a measurand specifies certain physical states and conditions.

EXAMPLE The velocity of sound in dry air of composition (mole fraction) $N_2 = 0{,}780\ 8$, $O_2 = 0{,}209\ 5$, Ar = 0,009 35, and $CO_2 = 0{,}000\ 35$ at the temperature $T = 273{,}15$ K and pressure $p = 101\ 325$ Pa.

D.2 The realized quantity

D.2.1 Ideally, the quantity realized for measurement would be fully consistent with the definition of the measurand. Often, however, such a quantity cannot be realized and the measurement is performed on a quantity that is an approximation of the measurand.

D.3 The "true" value and the corrected value

D.3.1 The result of the measurement of the realized quantity is corrected for the difference between that quantity and the measurand in order to predict what the measurement result would have been if the realized quantity had in fact fully satisfied the definition of the measurand. The result of the measurement of the realized quantity is also corrected for all other recognized significant systematic effects. Although the final corrected result is sometimes viewed as the best estimate of the "true" value of the measurand, in reality the result is simply the best estimate of the value of the quantity intended to be measured.

D.3.2 As an example, suppose that the measurand is the thickness of a given sheet of material at a specified temperature. The specimen is brought to a temperature near the specified temperature and its thickness at a particular place is measured with a micrometer. The thickness of the material at that place and temperature, under the pressure applied by the micrometer, is the realized quantity.

D.3.3 The temperature of the material at the time of the measurement and the applied pressure are determined. The uncorrected result of the measurement of the realized quantity is then corrected by taking into account the calibration curve of the micrometer, the departure of the temperature of the specimen from the specified temperature, and the slight compression of the specimen under the applied pressure.

D.3.4 The corrected result may be called the best estimate of the "true" value, "true" in the sense that it is the value of a quantity that is believed to satisfy fully the definition of the measurand; but had the micrometer been applied to a different part of the sheet of material, the realized quantity would have been different with a different "true" value. However, that "true" value would be consistent with the definition of the measurand because the latter did not specify that the thickness was to be determined at a particular place on the sheet. Thus in this case, because of an incomplete definition of the measurand, the "true" value has an uncertainty that can be evaluated from measurements made at different places on the sheet. At some level, every measurand has such an "intrinsic" uncertainty that can in principle be estimated in some way. This is the minimum uncertainty with which a measurand can be determined, and every measurement that achieves such an uncertainty may be viewed as the best possible measurement of the measurand. To obtain a value of the quantity in question having a smaller uncertainty requires that the measurand be more completely defined.

NOTE 1 In the example, the measurand's specification leaves many other matters in doubt that might conceivably affect the thickness: the barometric pressure, the humidity, the attitude of the sheet in the gravitational field, the way it is supported, etc.

NOTE 2 Although a measurand should be defined in sufficient detail that any uncertainty arising from its incomplete definition is negligible in comparison with the required accuracy of the measurement, it must be recognized that this may not always be practicable. The definition may, for example, be incomplete because it does not specify parameters that may have been assumed, unjustifiably, to have negligible effect; or it may imply conditions that can never be fully met and whose imperfect realization is difficult to take into account. For instance, in the example of D.1.2, the velocity of sound implies infinite plane waves of vanishingly small amplitude. To the extent that the measurement does not meet these conditions, diffraction and nonlinear effects need to be considered.

NOTE 3 Inadequate specification of the measurand can lead to discrepancies between the results of measurements of ostensibly the same quantity carried out in different laboratories.

D.3.5 The term "true value of a measurand" or of a quantity (often truncated to "true value") is avoided in this *Guide* because the word "true" is viewed as redundant. "Measurand" (see B.2.9) means "particular quantity subject to measurement", hence "value of a measurand" means "value of a particular quantity subject to measurement". Since "particular quantity" is generally understood to mean a definite or specified quantity (see B.2.1, Note 1), the adjective "true" in "true value of a measurand" (or in "true value of a quantity") is unnecessary — the "true" value of the measurand (or quantity) is simply the value of the measurand (or quantity). In addition, as indicated in the discussion above, a unique "true" value is only an idealized concept.

D.4 Error

A corrected measurement result is not the value of the measurand — that is, it is in error — because of imperfect measurement of the realized quantity due to random variations of the observations (random effects), inadequate determination of the corrections for systematic effects, and incomplete knowledge of certain physical phenomena (also systematic effects). Neither the value of the realized quantity nor the value of the measurand can ever be known exactly; all that can be known is their estimated values. In the example above, the measured thickness of the sheet *may* be in error, that is, may differ from the value of the measurand (the thickness of the sheet), because each of the following may combine to contribute an unknown error to the measurement result:

a) slight differences between the indications of the micrometer when it is repeatedly applied to the same realized quantity;

b) imperfect calibration of the micrometer;

c) imperfect measurement of the temperature and of the applied pressure;

d) incomplete knowledge of the effects of temperature, barometric pressure, and humidity on the specimen or the micrometer or both.

D.5 Uncertainty

D.5.1 Whereas the exact values of the contributions to the error of a result of a measurement are unknown and unknowable, the *uncertainties* associated with the random and systematic effects that give rise to the error can be evaluated. But, even if the evaluated uncertainties are small, there is still no guarantee that the error in the measurement result is small; for in the determination of a correction or in the assessment of incomplete knowledge, a systematic effect may have been overlooked because it is unrecognized. Thus the uncertainty of a result of a measurement is not necessarily an indication of the likelihood that the measurement result is near the value of the measurand; it is simply an estimate of the likelihood of nearness to the best value that is consistent with presently available knowledge.

D.5.2 Uncertainty of measurement is thus an expression of the fact that, for a given measurand and a given result of measurement of it, there is not one value but an infinite number of values dispersed about the result that are consistent with all of the observations and data and one's knowledge of the physical world, and that with varying degrees of credibility can be attributed to the measurand.

D.5.3 It is fortunate that in many practical measurement situations, much of the discussion of this annex does not apply. Examples are when the measurand is adequately well defined; when standards or instruments are calibrated using well-known reference standards that are traceable to national standards; and when the uncertainties of the calibration corrections are insignificant compared to the uncertainties arising from random effects on the indications of instruments, or from a limited number of observations (see E.4.3). Nevertheless, incomplete knowledge of influence quantities and their effects can often contribute significantly to the uncertainty of the result of a measurement.

D.6 Graphical representation

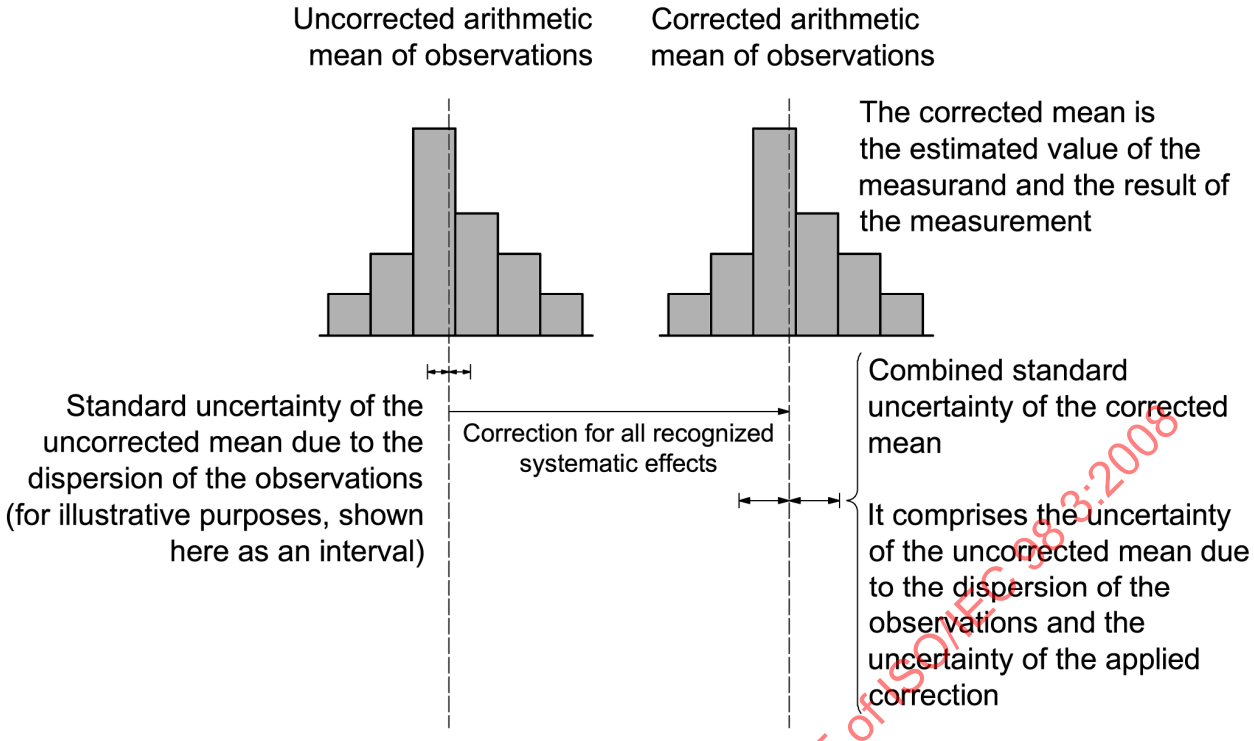
D.6.1 Figure D.1 depicts some of the ideas discussed in Clause 3 of this *Guide* and in this annex. It illustrates why the focus of this *Guide* is uncertainty and not error. The exact error of a result of a measurement is, in general, unknown and unknowable. All one can do is estimate the values of input quantities, including corrections for recognized systematic effects, together with their standard uncertainties (estimated standard deviations), either from unknown probability distributions that are sampled by means of repeated observations, or from subjective or *a priori* distributions based on the pool of available information; and then calculate the measurement result from the estimated values of the input quantities and the combined standard uncertainty of that result from the standard uncertainties of those estimated values. Only if there is a sound basis for believing that all of this has been done properly, with no significant systematic effects having been overlooked, can one assume that the measurement result is a reliable estimate of the value of the measurand and that its combined standard uncertainty is a reliable measure of its *possible* error.

NOTE 1 In Figure D.1 a), the observations are shown as a histogram for illustrative purposes [see 4.4.3 and Figure 1 b)].

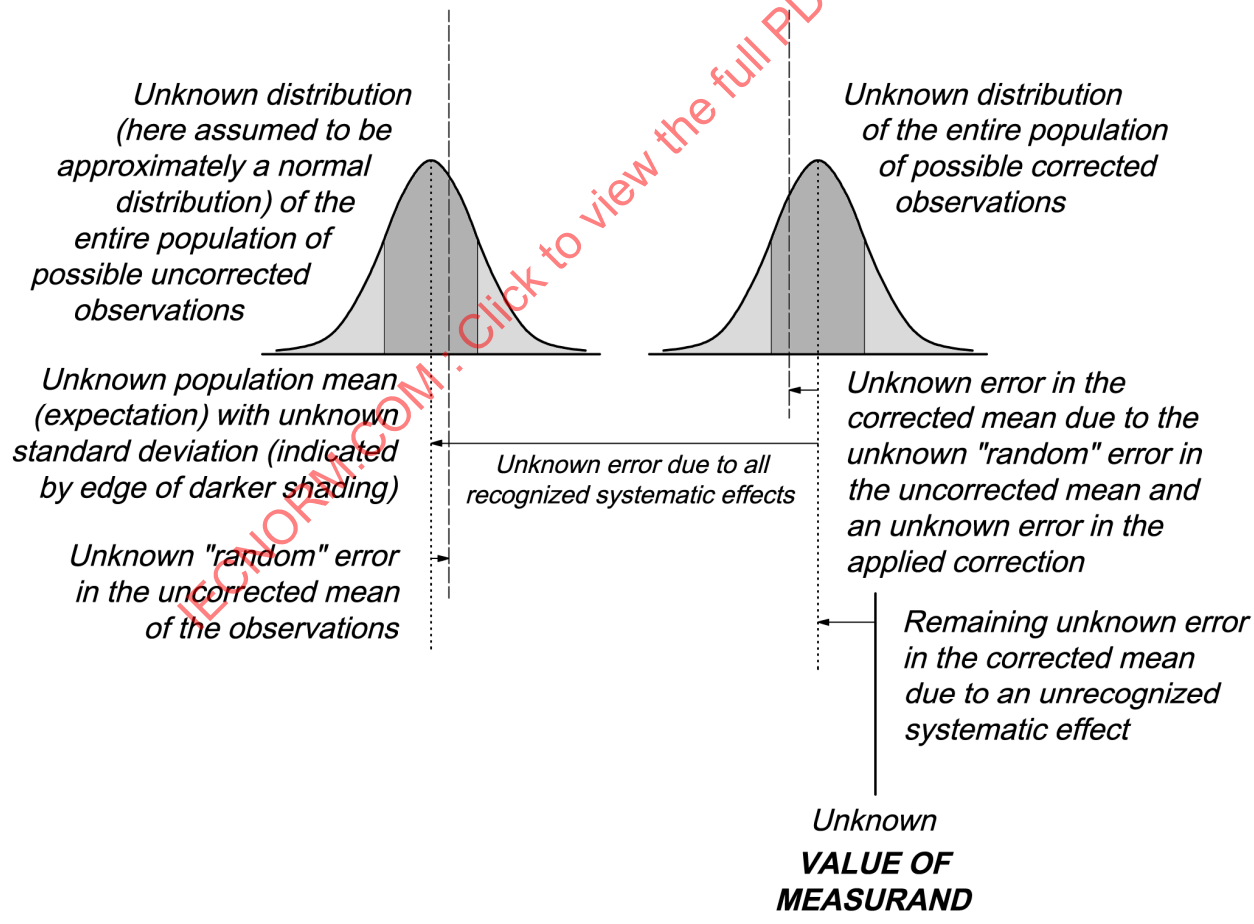
NOTE 2 The correction for an error is equal to the negative of the estimate of the error. Thus in Figure D.1, and in Figure D.2 as well, an arrow that illustrates the correction for an error is equal in length but points in the opposite direction to the arrow that would have illustrated the error itself, and vice versa. The text of the figure makes clear if a particular arrow illustrates a correction or an error.

D.6.2 Figure D.2 depicts some of the same ideas illustrated in Figure D.1 but in a different way. Moreover, it also depicts the idea that there can be many values of the measurand if the definition of the measurand is incomplete [entry *g)* of Figure D.2]. The uncertainty arising from this incompleteness of definition as measured by the variance is evaluated from measurements of multiple realizations of the measurand, using the same method, instruments, etc. (see D.3.4).

NOTE In the column headed “Variance”, the variances are understood to be the variances $u_i^2(y)$ defined in Equation (11a) in 5.1.3; hence they add linearly as shown.



a) Concepts based on observable quantities



b) *Ideal concepts based on unknowable quantities*

Figure D.1 — Graphical illustration of value, error, and uncertainty

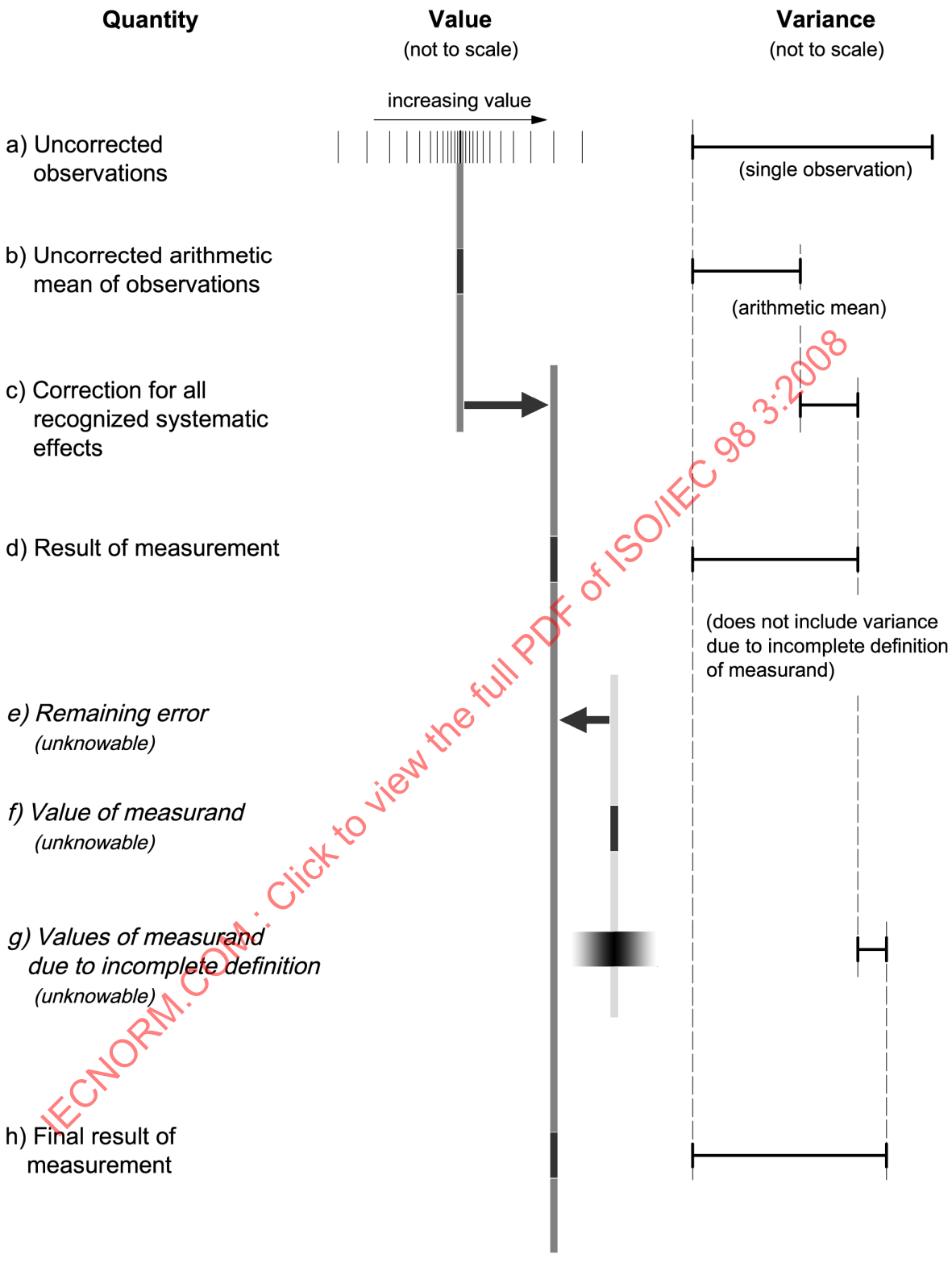


Figure D.2 — Graphical illustration of values, error, and uncertainty

Annex E

Motivation and basis for Recommendation INC-1 (1980)

This annex gives a brief discussion of both the motivation and statistical basis for Recommendation INC-1 (1980) of the Working Group on the Statement of Uncertainties upon which this *Guide* rests. For further discussion, see References [1, 2, 11, 12].

E.1 "Safe", "random", and "systematic"

E.1.1 This *Guide* presents a widely applicable method for evaluating and expressing uncertainty in measurement. It provides a realistic rather than a "safe" value of uncertainty based on the concept that there is no inherent difference between an uncertainty component arising from a random effect and one arising from a correction for a systematic effect (see 3.2.2 and 3.2.3). The method stands, therefore, in contrast to certain older methods that have the following two ideas in common.

E.1.2 The first idea is that the uncertainty reported should be "safe" or "conservative", meaning that it must never err on the side of being too small. In fact, because the evaluation of the uncertainty of a measurement result is problematic, it was often made deliberately large.

E.1.3 The second idea is that the influences that give rise to uncertainty were always recognizable as either "random" or "systematic" with the two being of different natures; the uncertainties associated with each were to be combined in their own way and were to be reported separately (or when a single number was required, combined in some specified way). In fact, the method of combining uncertainties was often designed to satisfy the safety requirement.

E.2 Justification for realistic uncertainty evaluations

E.2.1 When the value of a measurand is reported, the best estimate of its value and the best evaluation of the uncertainty of that estimate must be given, for if the uncertainty is to err, it is not normally possible to decide in which direction it should err "safely". An understatement of uncertainties might cause too much trust to be placed in the values reported, with sometimes embarrassing or even disastrous consequences. A deliberate overstatement of uncertainties could also have undesirable repercussions. It could cause users of measuring equipment to purchase instruments that are more expensive than they need, or it could cause costly products to be discarded unnecessarily or the services of a calibration laboratory to be rejected.

E.2.2 That is not to say that those using a measurement result could not apply their own multiplicative factor to its stated uncertainty in order to obtain an expanded uncertainty that defines an interval having a specified level of confidence and that satisfies their own needs, nor in certain circumstances that institutions providing measurement results could not routinely apply a factor that provides a similar expanded uncertainty that meets the needs of a particular class of users of their results. However, such factors (always to be stated) must be applied to the uncertainty as determined by a realistic method, and only *after* the uncertainty has been so determined, so that the interval defined by the expanded uncertainty has the level of confidence required and the operation may be easily reversed.

E.2.3 Those engaged in measurement often must incorporate in their analyses the results of measurements made by others, with each of these other results possessing an uncertainty of its own. In evaluating the uncertainty of their own measurement result, they need to have a best value, not a "safe" value, of the uncertainty of each of the results incorporated from elsewhere. Additionally, there must be a logical and simple way in which these imported uncertainties can be combined with the uncertainties of their own observations to give the uncertainty of their own result. Recommendation INC-1 (1980) provides such a way.

E.3 Justification for treating all uncertainty components identically

The focus of the discussion of this subclause is a simple example that illustrates how this *Guide* treats uncertainty components arising from random effects and from corrections for systematic effects in exactly the same way in the evaluation of the uncertainty of the result of a measurement. It thus exemplifies the viewpoint adopted in this *Guide* and cited in E.1.1, namely, that all components of uncertainty are of the same nature and are to be treated identically. The starting point of the discussion is a simplified derivation of the mathematical expression for the propagation of standard deviations, termed in this *Guide* the law of propagation of uncertainty.

E.3.1 Let the output quantity $z = f(w_1, w_2, ..., w_N)$ depend on N input quantities $w_1, w_2, ..., w_N$, where each w_i is described by an appropriate probability distribution. Expansion of f about the expectations of the w_i, $E(w_i) = \mu_i$, in a first-order Taylor series yields for small deviations of z about μ_z in terms of small deviations of w_i about μ_i,

$$z - \mu_z = \sum_{i=1}^{N} \frac{\partial f}{\partial w_i}\left(w_i - \mu_i\right) \tag{E.1}$$

where all higher-order terms are assumed to be negligible and $\mu_z = f(\mu_1, \mu_2, ..., \mu_N)$. The square of the deviation $z - \mu_z$ is then given by

$$\left(z - \mu_z\right)^2 = \left(\sum_{i=1}^{N} \frac{\partial f}{\partial w_i}\left(w_i - \mu_i\right)\right)^2 \tag{E.2a}$$

which may be written as

$$\left(z - \mu_z\right)^2 = \sum_{i=1}^{N}\left(\frac{\partial f}{\partial w_i}\right)^2 \left(w_i - \mu_i\right)^2 + 2\sum_{i=1}^{N-1}\sum_{j=i+1}^{N} \frac{\partial f}{\partial w_i}\frac{\partial f}{\partial w_j}\left(w_i - \mu_i\right)\left(w_j - \mu_j\right) \tag{E.2b}$$

The expectation of the squared deviation $(z - \mu_z)^2$ is the variance of z, that is, $E[(z - \mu_z)^2] = \sigma_z^2$, and thus Equation (E.2b) leads to

$$\sigma_z^2 = \sum_{i=1}^{N}\left(\frac{\partial f}{\partial w_i}\right)^2 \sigma_i^2 + 2\sum_{i=1}^{N-1}\sum_{j=i+1}^{N} \frac{\partial f}{\partial w_i}\frac{\partial f}{\partial w_j}\sigma_i \sigma_j r_{ij} \tag{E.3}$$

In this expression, $\sigma_i^2 = E[(w_i - \mu_i)^2]$ is the variance of w_i and $r_{ij} = \left(w_i, w_j\right) \Big/ \left(\sigma_i^2 \sigma_j^2\right)^{1/2}$ is the correlation coefficient of w_i and w_j, where $\left(w_i, w_j\right) = E[(w_i - \mu_i)(w_j - \mu_j)]$ is the covariance of w_i and w_j.

NOTE 1 σ_z^2 and σ_i^2 are, respectively, the central moments of order 2 (see C.2.13 and C.2.22) of the probability distributions of z and w_i. A probability distribution may be completely characterized by its expectation, variance, and higher-order central moments.

NOTE 2 Equation (13) in 5.2.2 [together with Equation (15)], which is used to calculate combined standard uncertainty, is identical to Equation (E.3) except that Equation (13) is expressed in terms of estimates of the variances, standard deviations, and correlation coefficients.

E.3.2 In the traditional terminology, Equation (E.3) is often called the "general law of error propagation", an appellation that is better applied to an expression of the form $\Delta z = \sum_{i=1}^{N} \left(\partial f / \partial w_i\right) \Delta w_i$, where Δz is the change in z due to (small) changes Δw_i in the w_i [see Equation (E.8)]. In fact, it is appropriate to call Equation (E.3) the law of propagation of uncertainty as is done in this *Guide* because it shows how the uncertainties of the input quantities w_i, taken equal to the standard deviations of the probability distributions of the w_i, combine to give the uncertainty of the output quantity z if that uncertainty is taken equal to the standard deviation of the probability distribution of z.

E.3.3 Equation (E.3) also applies to the propagation of multiples of standard deviations, for if each standard deviation σ_i is replaced by a multiple $k\sigma_i$, with the same k for each σ_i, the standard deviation of the output quantity z is replaced by $k\sigma_z$. However, it does not apply to the propagation of confidence intervals. If each σ_i is replaced with a quantity δ_i that defines an interval corresponding to a given level of confidence p, the resulting quantity for z, δ_z, will not define an interval corresponding to the same value of p unless all of the w_i are described by normal distributions. No such assumptions regarding the normality of the probability distributions of the quantities w_i are implied in Equation (E.3). More specifically, if in Equation (10) in 5.1.2 each standard uncertainty $u(x_i)$ is evaluated from independent repeated observations and multiplied by the t-factor appropriate for its degrees of freedom for a particular value of p (say $p = 95$ percent), the uncertainty of the estimate y will not define an interval corresponding to that value of p (see G.3 and G.4).

NOTE The requirement of normality when propagating confidence intervals using Equation (E.3) may be one of the reasons for the historic separation of the components of uncertainty derived from repeated observations, which were assumed to be normally distributed, from those that were evaluated simply as upper and lower bounds.

E.3.4 Consider the following example: z depends on only one input quantity w, $z = f(w)$, where w is estimated by averaging n values w_k of w; these n values are obtained from n independent repeated observations q_k of a random variable q; and w_k and q_k are related by

$$w_k = \alpha + \Downarrow q_k \tag{E.4}$$

Here α is a constant "systematic" offset or shift common to each observation, and $\Downarrow$ is a common scale factor. The offset and the scale factor, although fixed during the course of the observations, are assumed to be characterized by *a priori* probability distributions, with α and $\Downarrow$ the best estimates of the expectations of these distributions.

The best estimate of w is the arithmetic mean or average $\overline{w}$ obtained from

$$\overline{w} = \frac{1}{n}\sum_{k=1}^{n} w_k = \frac{1}{n}\sum_{k=1}^{n}(\alpha + \Downarrow q_k) \tag{E.5}$$

The quantity z is then estimated by $f(\overline{w}) = f(\alpha, \Downarrow, q_1, q_2, ..., q_n)$ and the estimate $u^2(z)$ of its variance $\sigma^2(z)$ is obtained from Equation (E.3). If for simplicity it is assumed that $z = w$ so that the best estimate of z is $z = f(\overline{w}) = \overline{w}$, then the estimate $u^2(z)$ can be readily found. Noting from Equation (E.5) that

$$\frac{\partial f}{\partial \alpha} = 1,$$

$$\frac{\partial f}{\partial \Downarrow} = \frac{1}{n}\sum_{k=1}^{n} q_k = \overline{q},$$

and

$$\frac{\partial f}{\partial q_k} = \frac{\Downarrow}{n},$$

denoting the estimated variances of α and $\Downarrow$ by $u^2(\alpha)$ and $u^2(\Downarrow)$, respectively, and assuming that the individual observations are uncorrelated, one finds from Equation (E.3)

$$u^2(z) = u^2(\alpha) + \overline{q}^2 u^2(\Downarrow) + \Downarrow^2 \frac{s^2(q_k)}{n} \tag{E.6}$$

where $s^2(q_k)$ is the experimental variance of the observations q_k calculated according to Equation (4) in 4.2.2, and $s^2(q_k)/n = s^2(\overline{q})$ is the experimental variance of the mean $\overline{q}$ [Equation (5) in 4.2.3].

E.3.5 In the traditional terminology, the third term on the right-hand side of Equation (E.6) is called a "random" contribution to the estimated variance $u^2(z)$ because it normally decreases as the number of observations n increases, while the first two terms are called "systematic" contributions because they do not depend on n.

Of more significance, in some traditional treatments of measurement uncertainty, Equation (E.6) is questioned because no distinction is made between uncertainties arising from systematic effects and those arising from random effects. In particular, combining variances obtained from *a priori* probability distributions with those obtained from frequency-based distributions is deprecated because the concept of probability is considered to be applicable *only* to events that can be repeated a large number of times under essentially the same conditions, with the probability p of an event ($0 \leqslant p \leqslant 1$) indicating the *relative frequency* with which the event will occur.

In contrast to this frequency-based point of view of probability, an equally valid viewpoint is that probability is a measure of the *degree of belief* that an event will occur [13, 14]. For example, suppose one has a chance of winning a small sum of money D and one is a rational bettor. One's degree of belief in event A occurring is $p = 0{,}5$ if one is indifferent to these two betting choices:

1) receiving D if event A occurs but nothing if it does not occur;

2) receiving D if event A does not occur but nothing if it does occur.

Recommendation INC-1 (1980) upon which this *Guide* rests implicitly adopts such a viewpoint of probability since it views expressions such as Equation (E.6) as the appropriate way to calculate the combined standard uncertainty of a result of a measurement.

E.3.6 There are three distinct advantages to adopting an interpretation of probability based on degree of belief, the standard deviation (standard uncertainty), and the law of propagation of uncertainty [Equation (E.3)] as the basis for evaluating and expressing uncertainty in measurement, as has been done in this *Guide*:

a) the law of propagation of uncertainty allows the combined standard uncertainty of one result to be readily incorporated in the evaluation of the combined standard uncertainty of another result in which the first is used;

b) the combined standard uncertainty can serve as the basis for calculating intervals that correspond in a realistic way to their required levels of confidence; and

c) it is unnecessary to classify components as "random" or "systematic" (or in any other manner) when evaluating uncertainty because all components of uncertainty are treated in the same way.

Benefit c) is highly advantageous because such categorization is frequently a source of confusion; an uncertainty component is not either "random" or "systematic". Its nature is conditioned by the use made of the corresponding quantity, or more formally, by the context in which the quantity appears in the mathematical model that describes the measurement. Thus, when its corresponding quantity is used in a different context, a "random" component may become a "systematic" component, and vice versa.

E.3.7 For the reason given in c) above, Recommendation INC-1 (1980) does not classify components of uncertainty as either "random" or "systematic". In fact, as far as the calculation of the combined standard uncertainty of a measurement result is concerned, there is no need to classify uncertainty components and thus no real need for any classificational scheme. Nonetheless, since convenient labels can sometimes be helpful in the communication and discussion of ideas, Recommendation INC-1 (1980) does provide a scheme for classifying the two distinct *methods* by which uncertainty components may be evaluated, "A" and "B" (see 0.7, 2.3.2, and 2.3.3).

Classifying the methods used to evaluate uncertainty components avoids the principal problem associated with classifying the components themselves, namely, the dependence of the classification of a component on how the corresponding quantity is used. However, classifying the methods rather than the components does not preclude gathering the individual components evaluated by the two methods into specific groups for a particular purpose in a given measurement, for example, when comparing the experimentally observed and theoretically predicted variability of the output values of a complex measurement system (see 3.4.3).

E.4 Standard deviations as measures of uncertainty

E.4.1 Equation (E.3) requires that no matter how the uncertainty of the estimate of an input quantity is obtained, it must be evaluated as a standard uncertainty, that is, as an estimated standard deviation. If some "safe" alternative is evaluated instead, it cannot be used in Equation (E.3). In particular, if the "maximum error bound" (the largest conceivable deviation from the putative best estimate) is used in Equation (E.3), the resulting uncertainty will have an ill-defined meaning and will be unusable by anyone wishing to incorporate it into subsequent calculations of the uncertainties of other quantities (see E.3.3).

E.4.2 When the standard uncertainty of an input quantity cannot be evaluated by an analysis of the results of an adequate number of repeated observations, a probability distribution must be adopted based on knowledge that is much less extensive than might be desirable. That does not, however, make the distribution invalid or unreal; like all probability distributions, it is an expression of what knowledge exists.

E.4.3 Evaluations based on repeated observations are not necessarily superior to those obtained by other means. Consider $s(\bar{q})$, the experimental standard deviation of the mean of n independent observations q_k of a normally distributed random variable q [see Equation (5) in 4.2.3]. The quantity $s(\bar{q})$ is a statistic (see C.2.23) that estimates $\sigma(\bar{q})$, the standard deviation of the probability distribution of $\bar{q}$, that is, the standard deviation of the distribution of the values of $\bar{q}$ that would be obtained if the measurement were repeated an infinite number of times. The variance $\sigma^2[s(\bar{q})]$ of $s(\bar{q})$ is given, approximately, by

$$\sigma^2\left[s(\bar{q})\right] \quad \sigma^2(\bar{q})/(2v) \tag{E.7}$$

where $v = n - 1$ is the degrees of freedom of $s(\bar{q})$ (see G.3.3). Thus the relative standard deviation of $s(\bar{q})$, which is given by the ratio $\sigma[s(\bar{q})]/\sigma(\bar{q})$ and which can be taken as a measure of the relative uncertainty of $s(\bar{q})$, is approximately $[2(n-1)]^{-1/2}$. This "uncertainty of the uncertainty" of $\bar{q}$, which arises from the purely statistical reason of limited sampling, can be surprisingly large; for $n = 10$ observations it is 24 percent. This and other values are given in Table E.1, which shows that the standard deviation of a statistically estimated standard deviation is not negligible for practical values of n. One may therefore conclude that Type A evaluations of standard uncertainty are not necessarily more reliable than Type B evaluations, and that in many practical measurement situations where the number of observations is limited, the components obtained from Type B evaluations may be better known than the components obtained from Type A evaluations.

E.4.4 It has been argued that, whereas the uncertainties associated with the application of a particular method of measurement are statistical parameters characterizing random variables, there are instances of a "truly systematic effect" whose uncertainty must be treated differently. An example is an offset having an unknown fixed value that is the same for every determination by the method due to a possible imperfection in the very principle of the method itself or one of its underlying assumptions. But if the possibility of such an offset is acknowledged to exist and its magnitude is believed to be possibly significant, then it can be described by a probability distribution, however simply constructed, based on the knowledge that led to the conclusion that it could exist and be significant. Thus, if one considers probability to be a measure of the degree of belief that an event will occur, the contribution of such a systematic effect can be included in the combined standard uncertainty of a measurement result by evaluating it as a standard uncertainty of an *a priori* probability distribution and treating it in the same manner as any other standard uncertainty of an input quantity.

EXAMPLE The specification of a particular measurement procedure requires that a certain input quantity be calculated from a specific power-series expansion whose higher-order terms are inexactly known. The systematic effect due to not being able to treat these terms exactly leads to an unknown fixed offset that cannot be experimentally sampled by repetitions of the procedure. Thus the uncertainty associated with the effect cannot be evaluated and included in the uncertainty of the final measurement result if a frequency-based interpretation of probability is strictly followed. However, interpreting probability on the basis of degree of belief allows the uncertainty characterizing the effect to be evaluated from an *a priori* probability distribution (derived from the available knowledge concerning the inexactly known terms) and to be included in the calculation of the combined standard uncertainty of the measurement result like any other uncertainty.

Table E.1 — $\sigma[s(\bar{q})]/\sigma(\bar{q})$, the standard deviation of the experimental standard deviation of the mean $\bar{q}$ of n independent observations of a normally distributed random variable q, relative to the standard deviation of that mean (a) (b)

Number of observations n	$\sigma[s(\bar{q})]/\sigma(\bar{q})$ (percent)
2	76
3	52
4	42
5	36
10	24
20	16
30	13
50	10

(a) The values given have been calculated from the exact expression for $\sigma[s(\bar{q})]/\sigma(\bar{q})$, not the approximate expression $[2(n-1)]^{-1/2}$.

(b) In the expression $\sigma[s(\bar{q})]/\sigma(\bar{q})$, the denominator $\sigma(\bar{q})$ is the expectation $E[S/\sqrt{n}]$ and the numerator $\sigma[s(\bar{q})]$ is the square root of the variance $V[S/\sqrt{n}]$, where S denotes a random variable equal to the standard deviation of n independent random variables $X_1, \ldots, X_n$, each having a normal distribution with mean value $\propto$ and variance σ^2:

$$S = \sqrt{\frac{1}{n-1}\sum_{i=1}^{n}\left(X_i - \bar{X}\right)^2}, \quad \bar{X} = \frac{1}{n}\sum_{i=1}^{n} X_i$$

The expectation and variance of S are given by:

$$E[S] = \sqrt{\frac{2}{n-1}}\,\frac{\Gamma(n/2)}{\Gamma\left[(n-1)/2\right]}\,\sigma, \quad V[S] = \sigma^2 - E[S]^2,$$

where $\Gamma(x)$ is the gamma function. Note that $E[S] < \sigma$ for a finite number n.

E.5 A comparison of two views of uncertainty

E.5.1 The focus of this *Guide* is on the measurement result and its evaluated uncertainty rather than on the unknowable quantities "true" value and error (see Annex D). By taking the operational views that the result of a measurement is simply the value attributed to the measurand and that the uncertainty of that result is a measure of the dispersion of the values that could reasonably be attributed to the measurand, this *Guide* in effect uncouples the often confusing connection between uncertainty and the unknowable quantities "true" value and error.

E.5.2 This connection may be understood by interpreting the derivation of Equation (E.3), the law of propagation of uncertainty, from the standpoint of "true" value and error. In this case, $\propto_i$ is viewed as the unknown, unique "true" value of input quantity w_i and each w_i is assumed to be related to its "true" value $\propto_i$ by $w_i = \propto_i + \varepsilon_i$, where ε_i is the error in w_i. The expectation of the probability distribution of each ε_i is assumed to be zero, $E(\varepsilon_i) = 0$, with variance $E(\varepsilon_i^2) = \sigma_i^2$. Equation (E.1) becomes then

$$\varepsilon_z = \sum_{i=1}^{N} \frac{\cdot f}{\cdot w_i}\varepsilon_i \tag{E.8}$$

where $\varepsilon_z = z - \propto_z$ is the error in z and $\propto_z$ is the "true" value of z. If one then takes the expectation of the square of ε_z, one obtains an equation identical in form to Equation (E.3) but in which $\sigma_z^2 = E(\varepsilon_z^2)$ is the variance of ε_z and $\cdot_{ij} =. (\varepsilon_i, \varepsilon_j)/(\sigma_i^2\sigma_j^2)^{1/2}$ is the correlation coefficient of ε_i and ε_j, where $\cdot(\varepsilon_i, \varepsilon_j) = E(\varepsilon_i\varepsilon_j)$ is the

covariance of ε_i and ε_j. The variances and correlation coefficients are thus associated with the *errors* of the input quantities rather than with the input quantities themselves.

NOTE It is assumed that probability is viewed as a measure of the degree of belief that an event will occur, implying that a systematic error may be treated in the same way as a random error and that ε_i represents either kind.

E.5.3 In practice, the difference in point of view does not lead to a difference in the numerical value of the measurement result or of the uncertainty assigned to that result.

First, in both cases, the best available estimates of the input quantities w_i are used to obtain the best estimate of z from the function f; it makes no difference *in the calculations* if the best estimates are viewed as the values most likely to be attributed to the quantities in question or the best estimates of their "true" values.

Second, because $\varepsilon_i = w_i - \infty_i$, and because the ∞_i represent unique, fixed values and hence have no uncertainty, the variances and standard deviations of the ε_i and w_i are identical. This means that in both cases, the standard uncertainties used as the estimates of the standard deviations σ_i to obtain the combined standard uncertainty of the measurement result are identical and will yield the same numerical value for that uncertainty. Again, it makes no difference *in the calculations* if a standard uncertainty is viewed as a measure of the dispersion of the probability distribution of an input quantity or as a measure of the dispersion of the probability distribution of the error of that quantity.

NOTE If the assumption of the note of E.5.2 had not been made, then the discussion of this subclause would not apply unless all of the estimates of the input quantities and the uncertainties of those estimates were obtained from the statistical analysis of repeated observations, that is, from Type A evaluations.

E.5.4 While the approach based on "true" value and error yields the same numerical results as the approach taken in this *Guide* (provided that the assumption of the note of E.5.2 is made), this *Guide*'s concept of uncertainty eliminates the confusion between error and uncertainty (see Annex D). Indeed, this *Guide*'s operational approach, wherein the focus is on the observed (or estimated) value of a quantity and the observed (or estimated) variability of that value, makes any mention of error entirely unnecessary.

Annex F

Practical guidance on evaluating uncertainty components

This annex gives additional suggestions for evaluating uncertainty components, mainly of a practical nature, that are intended to complement the suggestions already given in Clause 4.

F.1 Components evaluated from repeated observations: Type A evaluation of standard uncertainty

F.1.1 Randomness and repeated observations

F.1.1.1 Uncertainties determined from repeated observations are often contrasted with those evaluated by other means as being "objective", "statistically rigorous", etc. That incorrectly implies that they can be evaluated merely by the application of statistical formulae to the observations and that their evaluation does not require the application of some judgement.

F.1.1.2 It must first be asked, "To what extent are the repeated observations completely independent repetitions of the measurement procedure?" If all of the observations are on a single sample, and if sampling is part of the measurement procedure because the measurand is the property of a material (as opposed to the property of a given specimen of the material), then the observations have not been independently repeated; an evaluation of a component of variance arising from possible differences among samples must be added to the observed variance of the repeated observations made on the single sample.

If zeroing an instrument is part of the measurement procedure, the instrument ought to be rezeroed as part of every repetition, even if there is negligible drift during the period in which observations are made, for there is potentially a statistically determinable uncertainty attributable to zeroing.

Similarly, if a barometer has to be read, it should in principle be read for each repetition of the measurement (preferably after disturbing it and allowing it to return to equilibrium), for there may be a variation both in indication and in reading, even if the barometric pressure is constant.

F.1.1.3 Second, it must be asked whether all of the influences that are assumed to be random really are random. Are the means and variances of their distributions constant, or is there perhaps a drift in the value of an unmeasured influence quantity during the period of repeated observations? If there is a sufficient number of observations, the arithmetic means of the results of the first and second halves of the period and their experimental standard deviations may be calculated and the two means compared with each other in order to judge whether the difference between them is statistically significant and thus if there is an effect varying with time.

F.1.1.4 If the values of "common services" in the laboratory (electric-supply voltage and frequency, water pressure and temperature, nitrogen pressure, etc.) are influence quantities, there is normally a strongly nonrandom element in their variations that cannot be overlooked.

F.1.1.5 If the least significant figure of a digital indication varies continually during an observation due to "noise", it is sometimes difficult not to select unknowingly personally preferred values of that digit. It is better to arrange some means of freezing the indication at an arbitrary instant and recording the frozen result.

F.1.2 Correlations

Much of the discussion in this subclause is also applicable to Type B evaluations of standard uncertainty.

F.1.2.1 The covariance associated with the estimates of two input quantities X_i and X_j may be taken to be zero or treated as insignificant if

a) X_i and X_j are *uncorrelated* (the random variables, not the physical quantities that are assumed to be invariants — see 4.1.1, Note 1), for example, because they have been repeatedly but not simultaneously measured in *different* independent experiments or because they represent resultant quantities of *different* evaluations that have been made independently, or if

b) either of the quantities X_i or X_j can be treated as a constant, or if

c) there is insufficient information to evaluate the covariance associated with the estimates of X_i and X_j.

NOTE 1 On the other hand, in certain cases, such as the reference-resistance example of Note 1 to 5.2.2, it is apparent that the input quantities are fully correlated and that the standard uncertainties of their estimates combine linearly.

NOTE 2 Different experiments may not be independent if, for example, the same instrument is used in each (see F.1.2.3).

F.1.2.2 Whether or not two repeatedly and simultaneously observed input quantities are correlated may be determined by means of Equation (17) in 5.2.3. For example, if the frequency of an oscillator uncompensated or poorly compensated for temperature is an input quantity, if ambient temperature is also an input quantity, and if they are observed simultaneously, there may be a significant correlation revealed by the calculated covariance of the frequency of the oscillator and the ambient temperature.

F.1.2.3 In practice, input quantities are often correlated because the same physical measurement standard, measuring instrument, reference datum, or even measurement method having a significant uncertainty is used in the estimation of their values. Without loss of generality, suppose two input quantities X_1 and X_2 estimated by x_1 and x_2 depend on a set of uncorrelated variables $Q_1, Q_2, ..., Q_L$. Thus $X_1 = F(Q_1, Q_2, ..., Q_L)$ and $X_2 = G(Q_1, Q_2, ..., Q_L)$, although some of these variables may actually appear only in one function and not in the other. If $u^2(q_l)$ is the estimated variance associated with the estimate q_l of Q_l, then the estimated variance associated with x_1 is, from Equation (10) in 5.1.2,

$$u^2(x_1) = \sum_{l=1}^{L} \left(\frac{\partial F}{\partial q_l} \right)^2 u^2(q_l) \qquad \text{(F.1)}$$

with a similar expression for $u^2(x_2)$. The estimated covariance associated with x_1 and x_2 is given by

$$u(x_1, x_2) = \sum_{l=1}^{L} \frac{\partial F}{\partial q_l} \frac{\partial G}{\partial q_l} u^2(q_l) \qquad \text{(F.2)}$$

Because only those terms for which $\partial F/\partial q_l \neq 0$ and $\partial G/\partial q_l \neq 0$ for a given l contribute to the sum, the covariance is zero if no variable is common to both F and G.

The estimated correlation coefficient $r(x_1, x_2)$ associated with the two estimates x_1 and x_2 is determined from $u(x_1, x_2)$ [Equation (F.2)] and Equation (14) in 5.2.2, with $u(x_1)$ calculated from Equation (F.1) and $u(x_2)$ from a similar expression. [See also Equation (H.9) in H.2.3.] It is also possible for the estimated covariance associated with two input estimates to have both a statistical component [see Equation (17) in 5.2.3] and a component arising as discussed in this subclause.

EXAMPLE 1 A standard resistor R_S is used in the same measurement to determine both a current I and a temperature t. The current is determined by measuring, with a digital voltmeter, the potential difference across the terminals of the standard; the temperature is determined by measuring, with a resistance bridge and the standard, the resistance $R_t(t)$ of a calibrated resistive temperature sensor whose temperature-resistance relation in the range $15\ °C \leqslant t \leqslant 30\ °C$ is $t = aR_t^2(t) - t_0$, where a and t_0 are known constants. Thus the current is determined through the relation $I = V_S/R_S$ and the temperature through the relation $t = a\mathbb{R}^2(t)R_S^2 - t_0$, where $\mathbb{R}(t)$ is the measured ratio $R_t(t)/R_S$ provided by the bridge.

Since only the quantity R_S is common to the expression for I and t, Equation (F.2) yields for the covariance of I and t

$$u(I,t) = \frac{\partial I}{\partial R_S}\frac{\partial t}{\partial R_S}u^2(R_S) = \left(-\frac{V_S}{R_S^2}\right)\left[2a\mathcal{R}^2(t)R_S\right]u^2(R_S) = -\frac{2I(t+t_0)}{R_S^2}u^2(R_S)$$

(For simplicity of notation, in this example the same symbol is used for both the input quantity and its estimate.)

To obtain the numerical value of the covariance, one substitutes into this expression the numerical values of the measured quantities I and t, and the values of R_S and $u(R_S)$ given in the standard resistor's calibration certificate. The unit of $u(I, t)$ is clearly A·°C since the dimension of the relative variance $[u(R_S)/R_S]^2$ is one (that is, the latter is a so-called dimensionless quantity).

Further, let a quantity P be related to the input quantities I and t by $P = C_0 I^2/(T_0 + t)$, where C_0 and T_0 are known constants with negligible uncertainties [$u^2(C_0)$ 0, $u^2(T_0)$ 0]. Equation (13) in 5.2.2 then yields for the variance of P in terms of the variances of I and t and their covariance

$$\frac{u^2(P)}{P^2} = 4\frac{u^2(I)}{I^2} - 4\frac{u(I,t)}{I(T_0+t)} + \frac{u^2(t)}{(T_0+t)^2}$$

The variances $u^2(I)$ and $u^2(t)$ are obtained by the application of Equation (10) of 5.1.2 to the relations $I = V_S/R_S$ and $t = a\mathcal{R}^2(t)R_S^2 - t_0$. The results are

$$u^2(I)/I^2 = u^2(V_S)/V_S^2 + u^2(R_S)/R_S^2$$

$$u^2(t) = 4(t+t_0)^2 u^2(\mathcal{R})/\mathcal{R}^2 + 4(t+t_0)^2 u^2(R_S)/R_S^2$$

where for simplicity it is assumed that the uncertainties of the constants t_0 and a are also negligible. These expressions can be readily evaluated since $u^2(V_S)$ and $u^2(\mathcal{R})$ may be determined, respectively, from the repeated readings of the voltmeter and of the resistance bridge. Of course, any uncertainties inherent in the instruments themselves and in the measurement procedures employed must also be taken into account when $u^2(V_S)$ and $u^2(\mathcal{R})$ are determined.

EXAMPLE 2 In the example of Note 1 to 5.2.2, let the calibration of each resistor be represented by $R_i = \alpha_i R_S$, with $u(\alpha_i)$ the standard uncertainty of the measured ratio α_i as obtained from repeated observations. Further, let α_i 1 for each resistor, and let $u(\alpha_i)$ be essentially the same for each calibration so that $u(\alpha_i)$ $u(\alpha)$. Then Equations (F.1) and (F.2) yield $u^2(R_i) = R_S^2 u^2(\alpha) + u^2(R_S)$ and $u(R_i, R_j) = u^2(R_S)$. This implies through Equation (14) in 5.2.2 that the correlation coefficient of any two resistors (i . j) is

$$r(R_i, R_j) = r_{ij} = \left\{1 + \left[\frac{u(\alpha)}{u(R_S)/R_S}\right]^2\right\}^{-1}$$

Since $u(R_S)/R_S = 10^{-4}$, if $u(\alpha) = 100 \cdot 10^{-6}$, r_{ij} 0,5; if $u(\alpha) = 10 \cdot 10^{-6}$, r_{ij} 0,990; and if $u(\alpha) = 1 \cdot 10^{-6}$, r_{ij} 1,000. Thus as $u(\alpha) \to 0$, $r_{ij} \to 1$, and $u(R_i) \to u(R_S)$.

NOTE In general, in comparison calibrations such as this example, the estimated values of the calibrated items are correlated, with the degree of correlation depending upon the ratio of the uncertainty of the comparison to the uncertainty of the reference standard. When, as often occurs in practice, the uncertainty of the comparison is negligible with respect to the uncertainty of the standard, the correlation coefficients are equal to +1 and the uncertainty of each calibrated item is the same as that of the standard.

F.1.2.4 The need to introduce the covariance $u(x_i, x_j)$ can be bypassed if the original set of input quantities $X_1, X_2, \ldots, X_N$ upon which the measurand Y depends [see Equation (1) in 4.1] is redefined in such a way as to include as additional independent input quantities those quantities Q_l that are common to two or more of the original X_i. (It may be necessary to perform additional measurements to establish fully the relationship between Q_l and the affected X_i.) Nonetheless, in some situations it may be more convenient to retain covariances rather than to increase the number of input quantities. A similar process can be carried out on the observed covariances of simultaneous repeated observations [see Equation (17) in 5.2.3], but the identification of the appropriate additional input quantities is often *ad hoc* and nonphysical.

EXAMPLE If, in Example 1 of F.1.2.3, the expressions for I and t in terms of R_S are introduced into the expression for P, the result is

$$P = \frac{C_0 V_S^2}{R_S^2 \left[T_0 + a \not{l}^2 (t) R_S^2 - t_0 \right]}$$

and the correlation between I and t is avoided at the expense of replacing the input quantities I and t with the quantities V_S, R_S, and $\not{l}$. Since these quantities are uncorrelated, the variance of P can be obtained from Equation (10) in 5.1.2.

F.2 Components evaluated by other means: Type B evaluation of standard uncertainty

F.2.1 The need for Type B evaluations

If a measurement laboratory had limitless time and resources, it could conduct an exhaustive statistical investigation of every conceivable cause of uncertainty, for example, by using many different makes and kinds of instruments, different methods of measurement, different applications of the method, and different approximations in its theoretical models of the measurement. The uncertainties associated with all of these causes could then be evaluated by the statistical analysis of series of observations and the uncertainty of each cause would be characterized by a statistically evaluated standard deviation. In other words, all of the uncertainty components would be obtained from Type A evaluations. Since such an investigation is not an economic practicality, many uncertainty components must be evaluated by whatever other means is practical.

F.2.2 Mathematically determinate distributions

F.2.2.1 The resolution of a digital indication

One source of uncertainty of a digital instrument is the resolution of its indicating device. For example, even if the repeated indications were all identical, the uncertainty of the measurement attributable to repeatability would not be zero, for there is a range of input signals to the instrument spanning a known interval that would give the same indication. If the resolution of the indicating device is δx, the value of the stimulus that produces a given indication X can lie with equal probability anywhere in the interval $X - \delta x/2$ to $X + \delta x/2$. The stimulus is thus described by a rectangular probability distribution (see 4.3.7 and 4.4.5) of width δx with variance $u^2 = (\delta x)^2/12$, implying a standard uncertainty of $u = 0{,}29\delta x$ for any indication.

Thus a weighing instrument with an indicating device whose smallest significant digit is 1 g has a variance due to the resolution of the device of $u^2 = (1/12)\ \text{g}^2$ and a standard uncertainty of $u = \left(1/\sqrt{12}\right)\ \text{g} = 0{,}29\ \text{g}$.

F.2.2.2 Hysteresis

Certain kinds of hysteresis can cause a similar kind of uncertainty. The indication of an instrument may differ by a fixed and known amount according to whether successive readings are rising or falling. The prudent operator takes note of the direction of successive readings and makes the appropriate correction. But the direction of the hysteresis is not always observable: there may be hidden oscillations within the instrument about an equilibrium point so that the indication depends on the direction from which that point is finally approached. If the range of possible readings from that cause is δx, the variance is again $u^2 = (\delta x)^2/12$, and the standard uncertainty due to hysteresis is $u = 0{,}29\delta x$.

F.2.2.3 Finite-precision arithmetic

The rounding or truncation of numbers arising in automated data reduction by computer can also be a source of uncertainty. Consider, for example, a computer with a word length of 16 bits. If, in the course of computation, a number having this word length is subtracted from another from which it differs only in the 16th bit, only one significant bit remains. Such events can occur in the evaluation of “ill-conditioned” algorithms, and they can be difficult to predict. One may obtain an empirical determination of the uncertainty by increasing the most important input quantity to the calculation (there is frequently one that is proportional to

the magnitude of the output quantity) by small increments until the output quantity changes; the smallest change in the output quantity that can be obtained by such means may be taken as a measure of the uncertainty; if it is δx, the variance is $u^2 = (\delta x)^2/12$ and $u = 0{,}29\delta x$.

NOTE One may check the uncertainty evaluation by comparing the result of the computation carried out on the limited word-length machine with the result of the same computation carried out on a machine with a significantly larger word length.

F.2.3 Imported input values

F.2.3.1 An *imported* value for an input quantity is one that has not been estimated in the course of a given measurement but has been obtained elsewhere as the result of an independent evaluation. Frequently such an imported value is accompanied by some kind of statement about its uncertainty. For example, the uncertainty may be given as a standard deviation, a multiple of a standard deviation, or the half-width of an interval having a stated level of confidence. Alternatively, upper and lower bounds may be given, or no information may be provided about the uncertainty. In the latter case, those who use the value must employ their own knowledge about the likely magnitude of the uncertainty, given the nature of the quantity, the reliability of the source, the uncertainties obtained in practice for such quantities, etc.

NOTE The discussion of the uncertainty of imported input quantities is included in this subclause on Type B evaluation of standard uncertainty for convenience; the uncertainty of such a quantity could be composed of components obtained from Type A evaluations or components obtained from both Type A and Type B evaluations. Since it is unnecessary to distinguish between components evaluated by the two different methods in order to calculate a combined standard uncertainty, it is unnecessary to know the composition of the uncertainty of an imported quantity.

F.2.3.2 Some calibration laboratories have adopted the practice of expressing "uncertainty" in the form of upper and lower limits that define an interval having a "minimum" level of confidence, for example, "at least" 95 percent. This may be viewed as an example of a so-called "safe" uncertainty (see E.1.2), and it cannot be converted to a standard uncertainty without a knowledge of how it was calculated. If sufficient information is given, it may be recalculated in accordance with the rules of this *Guide*; otherwise an independent assessment of the uncertainty must be made by whatever means are available.

F.2.3.3 Some uncertainties are given simply as maximum bounds within which *all* values of the quantity are said to lie. It is a common practice to assume that all values within those bounds are equally probable (a rectangular probability distribution), but such a distribution should not be assumed if there is reason to expect that values within but close to the bounds are less likely than those nearer the centre of the bounds. A rectangular distribution of half-width a has a variance of $a^2/3$; a normal distribution for which a is the half-width of an interval having a level of confidence of 99,73 percent has a variance of $a^2/9$. It may be prudent to adopt a compromise between those values, for example, by assuming a triangular distribution for which the variance is $a^2/6$ (see 4.3.9 and 4.4.6).

F.2.4 Measured input values

F.2.4.1 Single observation, calibrated instruments

If an input estimate has been obtained from a single observation with a particular instrument that has been calibrated against a standard of small uncertainty, the uncertainty of the estimate is mainly one of repeatability. The variance of repeated measurements by the instrument may have been obtained on an earlier occasion, not necessarily at precisely the same value of the reading but near enough to be useful, and it may be possible to assume the variance to be applicable to the input value in question. If no such information is available, an estimate must be made based on the nature of the measuring apparatus or instrument, the known variances of other instruments of similar construction, etc.

F.2.4.2 Single observation, verified instruments

Not all measuring instruments are accompanied by a calibration certificate or a calibration curve. Most instruments, however, are constructed to a written standard and verified, either by the manufacturer or by an independent authority, to conform to that standard. Usually the standard contains metrological requirements, often in the form of "maximum permissible errors", to which the instrument is required to conform. The

compliance of the instrument with these requirements is determined by comparison with a reference instrument whose maximum allowed uncertainty is usually specified in the standard. This uncertainty is then a component of the uncertainty of the verified instrument.

If nothing is known about the characteristic error curve of the verified instrument it must be assumed that there is an equal probability that the error has any value within the permitted limits, that is, a rectangular probability distribution. However, certain types of instruments have characteristic curves such that the errors are, for example, likely always to be positive in part of the measuring range and negative in other parts. Sometimes such information can be deduced from a study of the written standard.

F.2.4.3 Controlled quantities

Measurements are frequently made under controlled reference conditions that are assumed to remain constant during the course of a series of measurements. For example, measurements may be performed on specimens in a stirred oil bath whose temperature is controlled by a thermostat. The temperature of the bath may be measured at the time of each measurement on a specimen, but if the temperature of the bath is cycling, the instantaneous temperature of the specimen may not be the temperature indicated by the thermometer in the bath. The calculation of the temperature fluctuations of the specimen based on heat-transfer theory, and of their variance, is beyond the scope of this *Guide*, but it must start from a known or assumed temperature cycle for the bath. That cycle may be observed by a fine thermocouple and a temperature recorder, but failing that, an approximation of it may be deduced from a knowledge of the nature of the controls.

F.2.4.4 Asymmetric distributions of possible values

There are occasions when all possible values of a quantity lie to one side of a single limiting value. For example, when measuring the fixed vertical height h (the measurand) of a column of liquid in a manometer, the axis of the height-measuring device may deviate from verticality by a small angle ε. The distance l determined by the device will always be *larger* than h; no values less than h are possible. This is because h is equal to the projection $l\cos\varepsilon$, implying $l = h/\cos\varepsilon$, and all values of $\cos\varepsilon$ are less than one; no values greater than one are possible. This so-called "cosine error" can also occur in such a way that the projection $h\cos\varepsilon$ of a measurand h is equal to the observed distance l, that is, $l = h\cos\varepsilon$, and the observed distance is always *less* than the measurand.

If a new variable $\delta = 1 - \cos\varepsilon$ is introduced, the two different situations are, assuming $\varepsilon \neq 0$ or $\delta \ll 1$ as is usually the case in practice,

$$h = \bar{l}\,(1-\delta) \tag{F.3a}$$

$$h = \bar{l}\,(1+\delta) \tag{F.3b}$$

Here $\bar{l}$, the best estimate of l, is the arithmetic mean or average of n independent repeated observations l_k of l with estimated variance $u^2(\bar{l})$ [see Equations (3) and (5) in 4.2]. Thus it follows from Equations (F.3a) and (F.3b) that to obtain an estimate of h or h requires an estimate of the correction factor δ, while to obtain the combined standard uncertainty of the estimate of h or h requires $u^2(\delta)$, the estimated variance of δ. More specifically, application of Equation (10) in 5.1.2 to Equations (F.3a) and (F.3b) yields for $u_c^2(h)$ and $u_c^2(h)$ (− and + signs, respectively)

$$u_c^2 = (1 \mp \delta)^2 u^2(\bar{l}) + \bar{l}^{\,2} u^2(\delta) \tag{F.4a}$$

$$u^2(\bar{l}) + \bar{l}^{\,2} u^2(\delta) \tag{F.4b}$$

To obtain estimates of the expected value of δ and the variance of δ, assume that the axis of the device used to measure the height of the column of liquid in the manometer is constrained to be fixed in a vertical plane and that the distribution of the values of the angle of inclination ε about its expected value of zero is a normal distribution with variance σ^2. Although ε can have both positive and negative values, $\delta = 1 - \cos\varepsilon$ is positive for all values of ε. If the misalignment of the axis of the device is assumed to be unconstrained, the orientation

of the axis can vary over a solid angle since it is capable of misalignment in azimuth as well, but φ is then always a positive angle.

In the constrained or one-dimensional case, the **probability element** $p(\varphi)\mathrm{d}\varphi$ (C.2.5, note) is proportional to $\{\exp[-\varphi^2/(2\sigma^2)]\}\mathrm{d}\varphi$; in the unconstrained or two-dimensional case, the probability element is proportional to $\{\exp[-\varphi^2/(2\sigma^2)]\}\sin\varphi\,\mathrm{d}\varphi$. The probability density functions $p(\delta)$ in the two cases are the expressions required to determine the expectation and variance of δ for use in Equations (F.3) and (F.4). They may readily be obtained from these probability elements because the angle φ may be assumed small, and hence $\delta = 1 - \cos\varphi$ and $\sin\varphi$ may be expanded to lowest order in φ. This yields $\delta \approx \varphi^2/2$, $\sin\varphi \approx \sqrt{2\delta}$, and $\mathrm{d}\varphi = \mathrm{d}\delta/\sqrt{2\delta}$. The probability density functions are then

$$p(\delta) = \frac{1}{\sigma\sqrt{\pi\delta}}\exp\left(-\delta/\sigma^2\right) \qquad \text{(F.5a)}$$

in one dimension

$$p(\delta) = \frac{1}{\sigma^2}\exp\left(-\delta/\sigma^2\right) \qquad \text{(F.5b)}$$

in two dimensions

where

$$\int_0^\infty p(\delta)\,\mathrm{d}\delta = 1$$

Equations (F.5a) and (F.5b), which show that the most probable value of the correction δ in both cases is zero, give in the one-dimensional case $E(\delta) = \sigma^2/2$ and $\mathrm{var}(\delta) = \sigma^4/2$ for the expectation and the variance of δ; and in the two-dimensional case $E(\delta) = \sigma^2$ and $\mathrm{var}(\delta) = \sigma^4$. Equations (F.3a), (F.3b), and (F.4b) become then

$$h = \bar{l}\left[1-(d/2)u^2(\varphi)\right] \qquad \text{(F.6a)}$$

$$h = \bar{l}\left[1+(d/2)u^2(\varphi)\right] \qquad \text{(F.6b)}$$

$$u_c^2(h) = u_c^2(h) = u^2(\bar{l}) + (d/2)\bar{l}^2u^4(\varphi) \qquad \text{(F.6c)}$$

where d is the dimensionality ($d = 1$ or 2) and $u(\varphi)$ is the standard uncertainty of the angle φ, taken to be the best estimate of the standard deviation σ of an assumed normal distribution and to be evaluated from all of the information available concerning the measurement (Type B evaluation). This is an example of a case where the estimate of the value of a measurand depends on the *uncertainty* of an input quantity.

Although Equations (F.6a) to (F.6c) are specific to the normal distribution, the analysis can be carried out assuming other distributions for φ. For example, if one assumes for φ a symmetric rectangular distribution with upper and lower bounds of $+\varphi_0$ and $-\varphi_0$ in the one-dimensional case and $+\varphi_0$ and zero in the two-dimensional case, $E(\delta) = \varphi_0^2/6$ and $\mathrm{var}(\delta) = \varphi_0^4/45$ in one dimension; and $E(\delta) = \varphi_0^2/4$ and $\mathrm{var}(\delta) = \varphi_0^4/48$ in two dimensions.

NOTE This is a situation where the expansion of the function $Y = f(X_1, X_2, \ldots, X_N)$ in a first-order Taylor series to obtain $u_c^2(y)$, Equation (10) in 5.1.2, is inadequate because of the nonlinearity of f: $\overline{\cos\varphi} \neq \cos\bar{\varphi}$ (see Note to 5.1.2, and H.2.4). Although the analysis can be carried out entirely in terms of φ, introducing the variable δ simplifies the problem.

Another example of a situation where all possible values of a quantity lie to one side of a single limiting value is the determination by titration of the concentration of a component in a solution where the end point is indicated by the triggering of a signal; the amount of reagent added is always more than that necessary to trigger the signal; it is never less. The excess titrated beyond the limit point is a required variable in the data

reduction, and the procedure in this (and in similar) cases is to assume an appropriate probability distribution for the excess and to use it to obtain the expected value of the excess and its variance.

EXAMPLE If a rectangular distribution of lower bound zero and upper bound C_0 is assumed for the excess z, then the expected value of the excess is $C_0/2$ with associated variance $C_0^2/12$. If the probability density function of the excess is taken as that of a normal distribution with $0 \leqslant z < 8$, that is, $p(z) = (\sigma\sqrt{\pi/2})^{-1} \exp\left[-z^2/(2\sigma^2)\right]$, then the expected value is $\sigma\sqrt{2/\pi}$ with variance $\sigma^2(1 - 2/\pi)$.

F.2.4.5 Uncertainty when corrections from a calibration curve are not applied

The note to 6.3.1 discusses the case where a known correction b for a significant systematic effect is not applied to the reported result of a measurement but instead is taken into account by enlarging the "uncertainty" assigned to the result. An example is replacement of an expanded uncertainty U with $U + b$, where U is an expanded uncertainty obtained under the assumption $b = 0$. This practice is sometimes followed in situations where all of the following conditions apply: the measurand Y is defined over a range of values of a parameter t, as in the case of a calibration curve for a temperature sensor; U and b also depend on t; and only a single value of "uncertainty" is to be given for all estimates $y(t)$ of the measurand over the range of possible values of t. In such situations the result of the measurement is often reported as $Y(t) = y(t) \pm [U_{\max} + b_{\max}]$, where the subscript "max" indicates that the maximum value of U and the maximum value of the known correction b over the range of values of t are used.

Although this *Guide* recommends that corrections be applied to measurement results for known significant systematic effects, this may not always be feasible in such a situation because of the unacceptable expense that would be incurred in calculating and applying an individual correction, and in calculating and using an individual uncertainty, for each value of $y(t)$.

A comparatively simple approach to this problem that is consistent with the principles of this *Guide* is as follows:

Compute a *single* mean correction $\bar{b}$ from

$$\bar{b} = \frac{1}{t_2 - t_1} \int_{t_1}^{t_2} b(t)\,\mathrm{d}t \tag{F.7a}$$

where t_1 and t_2 define the range of interest of the parameter t, and take the best estimate of $Y(t)$ to be $y_3(t) = y(t) + \bar{b}$, where $y(t)$ is the best uncorrected estimate of $Y(t)$. The variance associated with the mean correction $\bar{b}$ over the range of interest is given by

$$u^2(\bar{b}) = \frac{1}{t_2 - t_1} \int_{t_1}^{t_2} \left[b(t) - \bar{b}\right]^2 \mathrm{d}t \tag{F.7b}$$

not taking into account the uncertainty of the actual determination of the correction $b(t)$. The mean variance of the correction $b(t)$ due to its actual determination is given by

$$\overline{u^2\left[b(t)\right]} = \frac{1}{t_2 - t_1} \int_{t_1}^{t_2} u^2\left[b(t)\right] \mathrm{d}t \tag{F.7c}$$

where $u^2[b(t)]$ is the variance of the correction $b(t)$. Similarly, the mean variance of $y(t)$ arising from all sources of uncertainty other than the correction $b(t)$ is obtained from

$$\overline{u^2\left[y(t)\right]} = \frac{1}{t_2 - t_1} \int_{t_1}^{t_2} u^2\left[y(t)\right] \mathrm{d}t \tag{F.7d}$$

where $u^2[y(t)]$ is the variance of $y(t)$ due to all uncertainty sources other than $b(t)$. The single value of standard uncertainty to be used for *all* estimates $y_3(t) = y(t) + \bar{b}$ of the measurand $Y(t)$ is then the positive square root of

$$u_c^2(y_3) = \overline{u^2\left[y(t)\right]} + \overline{u^2\left[b(t)\right]} + u^2(\bar{b}) \tag{F.7e}$$

An expanded uncertainty U may be obtained by multiplying $u_c(y)$ by an appropriately chosen coverage factor k, $U = ku_c(y)$, yielding $Y(t) = y(t) \pm U = y(t) + \bar{b} \pm U$. However, the use of the same average correction for all values of t rather than the correction appropriate for each value of t must be recognized and a clear statement given as to what U represents.

F.2.5 Uncertainty of the method of measurement

F.2.5.1 Perhaps the most difficult uncertainty component to evaluate is that associated with the method of measurement, especially if the application of that method has been shown to give results with less variability than those of any other that is known. But it is likely that there are other methods, some of them as yet unknown or in some way impractical, that would give systematically different results of apparently equal validity. This implies an *a priori* probability distribution, not a distribution from which samples can be readily drawn and treated statistically. Thus, even though the uncertainty of the method may be the dominant one, the only information often available for evaluating its standard uncertainty is one's existing knowledge of the physical world. (See also E.4.4.)

NOTE Determining the same measurand by different methods, either in the same laboratory or in different laboratories, or by the same method in different laboratories, can often provide valuable information about the uncertainty attributable to a particular method. In general, the exchange of measurement standards or reference materials between laboratories for independent measurement is a useful way of assessing the reliability of evaluations of uncertainty and of identifying previously unrecognized systematic effects.

F.2.6 Uncertainty of the sample

F.2.6.1 Many measurements involve comparing an unknown object with a known standard having similar characteristics in order to calibrate the unknown. Examples include end gauges, certain thermometers, sets of masses, resistors, and high purity materials. In most such cases, the measurement methods are not especially sensitive to, or adversely affected by, sample selection (that is, the particular unknown being calibrated), sample treatment, or the effects of various environmental influence quantities because the unknown and standard respond in generally the same (and often predictable) way to such variables.

F.2.6.2 In some practical measurement situations, sampling and specimen treatment play a much larger role. This is often the case for the chemical analysis of natural materials. Unlike man-made materials, which may have proven homogeneity to a level beyond that required for the measurement, natural materials are often very inhomogeneous. This inhomogeneity leads to two additional uncertainty components. Evaluation of the first requires determining how adequately the sample selected represents the parent material being analysed. Evaluation of the second requires determining the extent to which the secondary (unanalysed) constituents influence the measurement and how adequately they are treated by the measurement method.

F.2.6.3 In some cases, careful design of the experiment may make it possible to evaluate statistically the uncertainty due to the sample (see H.5 and H.5.3.2). Usually, however, especially when the effects of environmental influence quantities on the sample are significant, the skill and knowledge of the analyst derived from experience and all of the currently available information are required for evaluating the uncertainty.

Annex G

Degrees of freedom and levels of confidence

G.1 Introduction

G.1.1 This annex addresses the general question of obtaining from the estimate y of the measurand Y, and from the combined standard uncertainty $u_c(y)$ of that estimate, an expanded uncertainty $U_p = k_p u_c(y)$ that defines an interval $y - U_p \leqslant Y \leqslant y + U_p$ that has a high, specified coverage probability or level of confidence p. It thus deals with the issue of determining the coverage factor k_p that produces an interval about the measurement result y that may be expected to encompass a large, specified fraction p of the distribution of values that could reasonably be attributed to the measurand Y (see Clause 6).

G.1.2 In most practical measurement situations, the calculation of intervals having specified levels of confidence — indeed, the estimation of most individual uncertainty components in such situations — is at best only approximate. Even the experimental standard deviation of the mean of as many as 30 repeated observations of a quantity described by a normal distribution has itself an uncertainty of about 13 percent (see Table E.1 in Annex E).

In most cases, it does not make sense to try to distinguish between, for example, an interval having a level of confidence of 95 percent (one chance in 20 that the value of the measurand Y lies outside the interval) and either a 94 percent or 96 percent interval (1 chance in 17 and 25, respectively). Obtaining justifiable intervals with levels of confidence of 99 percent (1 chance in 100) and higher is especially difficult, even if it is assumed that no systematic effects have been overlooked, because so little information is generally available about the most extreme portions or "tails" of the probability distributions of the input quantities.

G.1.3 To obtain the value of the coverage factor k_p that produces an interval corresponding to a specified level of confidence p requires detailed knowledge of the probability distribution characterized by the measurement result and its combined standard uncertainty. For example, for a quantity z described by a normal distribution with expectation $\propto_z$ and standard deviation σ, the value of k_p that produces an interval $\propto_z \pm k_p\sigma$ that encompasses the fraction p of the distribution, and thus has a coverage probability or level of confidence p, can be readily calculated. Some examples are given in Table G.1.

Table G.1 — Value of the coverage factor k_p that produces an interval having level of confidence p assuming a normal distribution

Level of confidence p (percent)	Coverage factor k_p
68,27	1
90	1,645
95	1,960
95,45	2
99	2,576
99,73	3

NOTE By contrast, if z is described by a rectangular probability distribution with expectation $\propto_z$ and standard deviation $\sigma = a/\sqrt{3}$, where a is the half-width of the distribution, the level of confidence p is 57,74 percent for $k_p = 1$; 95 percent for $k_p = 1{,}65$; 99 percent for $k_p = 1{,}71$; and 100 percent for $k_p \geqslant \sqrt{3}$ 1,73; the rectangular distribution is "narrower" than the normal distribution in the sense that it is of finite extent and has no "tails".

G.1.4 If the probability distributions of the input quantities $X_1, X_2, ..., X_N$ upon which the measurand Y depends are known [their expectations, variances, and higher moments (see C.2.13 and C.2.22) if the distributions are not normal distributions], and if Y is a linear function of the input quantities, $Y = c_1X_1 + c_2X_2 + ... + c_NX_N$, then the probability distribution of Y may be obtained by convolving the individual probability distributions [10]. Values of k_p that produce intervals corresponding to specified levels of confidence p may then be calculated from the resulting convolved distribution.

G.1.5 If the functional relationship between Y and its input quantities is nonlinear and a first-order Taylor series expansion of the relationship is not an acceptable approximation (see 5.1.2 and 5.1.5), then the probability distribution of Y cannot be obtained by convolving the distributions of the input quantities. In such cases, other analytical or numerical methods are required.

G.1.6 In practice, because the parameters characterizing the probability distributions of input quantities are usually estimates, because it is unrealistic to expect that the level of confidence to be associated with a given interval can be known with a great deal of exactness, and because of the complexity of convolving probability distributions, such convolutions are rarely, if ever, implemented when intervals having specified levels of confidence need to be calculated. Instead, approximations are used that take advantage of the Central Limit Theorem.

G.2 Central Limit Theorem

G.2.1 If $Y = c_1X_1 + c_2X_2 + ... + c_NX_N = \sum_{i=1}^{N} c_iX_i$ and all the X_i are characterized by normal distributions, then the resulting convolved distribution of Y will also be normal. However, even if the distributions of the X_i are not normal, the distribution of Y may often be approximated by a normal distribution because of the Central Limit Theorem. This theorem states that the distribution of Y will be *approximately normal* with expectation $E(Y) = \sum_{i=1}^{N} c_i E(X_i)$ and variance $\sigma^2(Y) = \sum_{i=1}^{N} c_i^2\sigma^2(X_i)$, where $E(X_i)$ is the expectation of X_i and $\sigma^2(X_i)$ is the variance of X_i, if the X_i are independent and $\sigma^2(Y)$ is much larger than any single component $c_i^2\sigma^2(X_i)$ from a non-normally distributed X_i.

G.2.2 The Central Limit Theorem is significant because it shows the very important role played by the variances of the probability distributions of the input quantities, compared with that played by the higher moments of the distributions, in determining the form of the resulting convolved distribution of Y. Further, it implies that the convolved distribution converges towards the normal distribution as the number of input quantities contributing to $\sigma^2(Y)$ increases; that the convergence will be more rapid the closer the values of $c_i^2\sigma^2(X_i)$ are to each other (equivalent in practice to each input estimate x_i contributing a comparable uncertainty to the uncertainty of the estimate y of the measurand Y); and that the closer the distributions of the X_i are to being normal, the fewer X_i are required to yield a normal distribution for Y.

EXAMPLE The rectangular distribution (see 4.3.7 and 4.4.5) is an extreme example of a non-normal distribution, but the convolution of even as few as *three* such distributions of equal width is approximately normal. If the half-width of each of the three rectangular distributions is a so that the variance of each is $a^2/3$, the variance of the convolved distribution is $\sigma^2 = a^2$. The 95 percent and 99 percent intervals of the convolved distribution are defined by $1{,}937\sigma$ and $2{,}379\sigma$, respectively, while the corresponding intervals for a normal distribution with the same standard deviation σ are defined by $1{,}960\sigma$ and $2{,}576\sigma$ (see Table G.1) [10].

NOTE 1 For every interval with a level of confidence p greater than about 91,7 percent, the value of k_p for a normal distribution is larger than the corresponding value for the distribution resulting from the convolution of any number and size of rectangular distributions.

NOTE 2 It follows from the Central Limit Theorem that the probability distribution of the arithmetic mean $\bar{q}$ of n observations q_k of a random variable q with expectation $\propto_q$ and finite standard deviation σ approaches a normal distribution with mean $\propto_q$ and standard deviation $\sigma/\sqrt{n}$ as $n \to 8$, whatever may be the probability distribution of q.

G.2.3 A practical consequence of the Central Limit Theorem is that when it can be established that its requirements are approximately met, in particular, if the combined standard uncertainty $u_c(y)$ is not dominated by a standard uncertainty component obtained from a Type A evaluation based on just a few observations, or by a standard uncertainty component obtained from a Type B evaluation based on an assumed rectangular

distribution, a reasonable first approximation to calculating an expanded uncertainty $U_p = k_p u_c(y)$ that provides an interval with level of confidence p is to use for k_p a value from the normal distribution. The values most commonly used for this purpose are given in Table G.1.

G.3 The t-distribution and degrees of freedom

G.3.1 To obtain a better approximation than simply using a value of k_p from the normal distribution as in G.2.3, it must be recognized that the calculation of an interval having a specified level of confidence requires, not the distribution of the variable $[Y - E(Y)]/\sigma(Y)$, but the distribution of the variable $(y - Y)/u_c(y)$. This is because in practice, all that is usually available are y, the estimate of Y as obtained from $y = \sum_{i=1}^{N} c_i x_i$, where x_i is the estimate of X_i; and the combined variance associated with y, $u_c^2(y)$, evaluated from $u_c^2(y) = \sum_{i=1}^{N} c_i^2 u^2(x_i)$, where $u(x_i)$ is the standard uncertainty (estimated standard deviation) of the estimate x_i.

NOTE Strictly speaking, in the expression $(y - Y)/u_c(y)$, Y should read $E(Y)$. For simplicity, such a distinction has only been made in a few places in this *Guide*. In general, the same symbol has been used for the physical quantity, the random variable that represents that quantity, and the expectation of that variable (see 4.1.1, notes).

G.3.2 If z is a normally distributed random variable with expectation $\propto_z$ and standard deviation σ, and $\bar{z}$ is the arithmetic mean of n independent observations z_k of z with $s(\bar{z})$ the experimental standard deviation of $\bar{z}$ [see Equations (3) and (5) in 4.2], then the distribution of the variable $t = (\bar{z} - \propto_z)/s(\bar{z})$ is the **t-distribution** or **Student's distribution** (C.3.8) with $\nu = n - 1$ degrees of freedom.

Consequently, if the measurand Y is simply a single normally distributed quantity X, $Y = X$; and if X is estimated by the arithmetic mean $\bar{X}$ of n independent repeated observations X_k of X, with experimental standard deviation of the mean $s(\bar{X})$, then the best estimate of Y is $y = \bar{X}$ and the experimental standard deviation of that estimate is $u_c(y) = s(\bar{X})$. Then $t = (\bar{z} - \propto_z)/s(\bar{z}) = (\bar{X} - X)/s(\bar{X}) = (y - Y)/u_c(y)$ is distributed according to the t-distribution with

$$\Pr\left[-t_p(\nu) \leqslant t \leqslant t_p(\nu)\right] = p \tag{G.1a}$$

or

$$\Pr\left[-t_p(\nu) \leqslant (y - Y)/u_c(y) \leqslant t_p(\nu)\right] = p \tag{G.1b}$$

which can be rewritten as

$$\Pr\left[y - t_p(\nu)\, u_c(y) \leqslant Y \leqslant y + t_p(\nu)\, u_c(y)\right] = p \tag{G,1c}$$

In these expressions, Pr[ ] means "probability of" and the t-factor $t_p(\nu)$ is the value of t for a given value of the parameter ν — the degrees of freedom (see G.3.3) — such that the fraction p of the t distribution is encompassed by the interval $-t_p(\nu)$ to $+t_p(\nu)$. Thus the expanded uncertainty

$$U_p = k_p u_c(y) = t_p(\nu) u_c(y) \tag{G.1d}$$

defines an interval $y - U_p$ to $y + U_p$, conveniently written as $Y = y \pm U_p$, that may be expected to encompass a fraction p of the distribution of values that could reasonably be attributed to Y, and p is the coverage probability or level of confidence of the interval.

G.3.3 The degrees of freedom ν is equal to $n - 1$ for a single quantity estimated by the arithmetic mean of n independent observations, as in G.3.2. If n independent observations are used to determine both the slope and intercept of a straight line by the method of least squares, the degrees of freedom of their respective standard uncertainties is $\nu = n - 2$. For a least-squares fit of m parameters to n data points, the degrees of freedom of the standard uncertainty of each parameter is $\nu = n - m$. (See Reference [15] for a further discussion of degrees of freedom.)

G.3.4 Selected values of $t_p(\nu)$ for different values of ν and various values of p are given in Table G.2 at the end of this annex. As $\nu \to \infty$ the t-distribution approaches the normal distribution and $t_p(\nu) \quad (1 + 2/\nu)^{1/2} k_p$, where in this expression k_p is the coverage factor required to obtain an interval with level of confidence p for a normally distributed variable. Thus the value of $t_p(\infty)$ in Table G.2 for a given p equals the value of k_p in Table G.1 for the same p.

NOTE Often, the t-distribution is tabulated in quantiles; that is, values of the quantile $t_{1-\alpha}$ are given, where $1 - \alpha$ denotes the cumulative probability and the relation

$$1-\alpha = \int_{-\infty}^{t_{1-\alpha}} f(t,\nu)\,\mathrm{d}t$$

defines the quantile, where f is the probability density function of t. Thus t_p and $t_{1-\alpha}$ are related by $p = 1 - 2\alpha$. For example, the value of the quantile $t_{0,975}$, for which $1 - \alpha = 0{,}975$ and $\alpha = 0{,}025$, is the same as $t_p(\nu)$ for $p = 0{,}95$.

G.4 Effective degrees of freedom

G.4.1 In general, the t-distribution will not describe the distribution of the variable $(y - Y)/u_c(y)$ if $u_c^2(y)$ is the sum of two or more estimated variance components $u_i^2(y) = c_i^2 u^2(x_i)$ (see 5.1.3), even if each x_i is the estimate of a normally distributed input quantity X_i. However, the distribution of that variable may be approximated by a t-distribution with an *effective* degrees of freedom ν_{eff} obtained from the Welch-Satterthwaite formula [16], [17], [18]

$$\frac{u_c^4(y)}{\nu_{\text{eff}}} = \sum_{i=1}^{N} \frac{u_i^4(y)}{\nu_i} \tag{G.2a}$$

or

$$\nu_{\text{eff}} = \frac{u_c^4(y)}{\sum_{i=1}^{N} \dfrac{u_i^4(y)}{\nu_i}} \tag{G.2b}$$

with

$$\nu_{\text{eff}} \leqslant \sum_{i=1}^{N} \nu_i \tag{G.2c}$$

where $u_c^2(y) = \sum_{i=1}^{N} u_i^2(y)$ (see 5.1.3). The expanded uncertainty $U_p = k_p u_c(y) = t_p(\nu_{\text{eff}}) u_c(y)$ then provides an interval $Y = y \pm U_p$ having an approximate level of confidence p.

NOTE 1 If the value of ν_{eff} obtained from Equation (G.2b) is not an integer, which will usually be the case in practice, the corresponding value of t_p may be found from Table G.2 by interpolation or by truncating ν_{eff} to the next lower integer.

NOTE 2 If an input estimate x_i is itself obtained from two or more other estimates, then the value of ν_i to be used with $u_i^4(y) = [c_i^2 u^2(x_i)]^2$ in the denominator of Equation (G.2b) is the effective degrees of freedom calculated from an expression equivalent to Equation (G.2b).

NOTE 3 Depending upon the needs of the potential users of a measurement result, it may be useful, in addition to ν_{eff}, to calculate and report also values for ν_{effA} and ν_{effB}, computed from Equation (G.2b) treating separately the standard uncertainties obtained from Type A and Type B evaluations. If the contributions to $u_c^2(y)$ of the Type A and Type B standard uncertainties alone are denoted, respectively, by $u_{cA}^2(y)$ and $u_{cB}^2(y)$, the various quantities are related by

$$u_c^2(y) = u_{cA}^2(y) + u_{cB}^2(y)$$

$$\frac{u_c^4(y)}{\nu_{\text{eff}}} = \frac{u_{cA}^4(y)}{\nu_{\text{effA}}} + \frac{u_{cB}^4(y)}{\nu_{\text{effB}}}$$

EXAMPLE Consider that $Y = f(X_1, X_2, X_3) = bX_1X_2X_3$ and that the estimates x_1, x_2, x_3 of the normally distributed input quantities X_1, X_2, X_3 are the arithmetic means of $n_1 = 10$, $n_2 = 5$, and $n_3 = 15$ independent repeated observations, respectively, with relative standard uncertainties $u(x_1)/x_1 = 0{,}25$ percent, $u(x_2)/x_2 = 0{,}57$ percent, and $u(x_3)/x_3 = 0{,}82$ percent. In this case, $c_i = \partial f/\partial X_i = Y/X_i$ (to be evaluated at x_1, x_2, x_3 — see 5.1.3, Note 1), $[u_c(y)/y]^2 = \sum_{i=1}^{3}[u(x_i)/x_i]^2 = (1{,}03 \text{ percent})^2$ (see Note 2 to 5.1.6), and Equation (G.2b) becomes

$$\nu_{\text{eff}} = \frac{\left[u_c(y)/y\right]^4}{\sum_{i=1}^{3}\frac{\left[u(x_i)/x_i\right]^4}{\nu_i}}$$

Thus

$$\nu_{\text{eff}} = \frac{1{,}03^4}{\frac{0{,}25^4}{10-1} + \frac{0{,}57^4}{5-1} + \frac{0{,}82^4}{15-1}} = 19{,}0$$

The value of t_p for $p = 95$ percent and $\nu = 19$ is, from Table G.2, $t_{95}(19) = 2{,}09$; hence the relative expanded uncertainty for this level of confidence is $U_{95} = 2{,}09 \cdot (1{,}03 \text{ percent}) = 2{,}2$ percent. It may then be stated that $Y = y \pm U_{95} = y(1 \pm 0{,}022)$ (y to be determined from $y = bx_1x_2x_3$), or that $0{,}978y \leqslant Y \leqslant 1{,}022y$, and that the level of confidence to be associated with the interval is approximately 95 percent.

G.4.2 In practice, $u_c(y)$ depends on standard uncertainties $u(x_i)$ of input estimates of both normally and non-normally distributed input quantities, and the $u(x_i)$ are obtained from both frequency-based and *a priori* probability distributions (that is, from both Type A and Type B evaluations). A similar statement applies to the estimate y and input estimates x_i upon which y depends. Nevertheless, the probability distribution of the function $t = (y - Y)/u_c(y)$ can be approximated by the t-distribution if it is expanded in a Taylor series about its expectation. In essence, this is what is achieved, in the lowest order approximation, by the Welch-Satterthwaite formula, Equation (G.2a) or Equation (G.2b).

The question arises as to the degrees of freedom to assign to a standard uncertainty obtained from a Type B evaluation when ν_{eff} is calculated from Equation (G.2b). Since the appropriate definition of degrees of freedom recognizes that ν as it appears in the t-distribution is a measure of the uncertainty of the variance $s^2(\bar{z})$, Equation (E.7) in E.4.3 may be used to define the degrees of freedom ν_i,

$$\nu_i \quad \frac{1}{2}\frac{u^2(x_i)}{\sigma^2\left[u(x_i)\right]} \quad \frac{1}{2}\left[\frac{\cdot\, u(x_i)}{u(x_i)}\right]^{-2} \tag{G.3}$$

The quantity in large brackets is the relative uncertainty of $u(x_i)$; for a Type B evaluation of standard uncertainty it is a subjective quantity whose value is obtained by scientific judgement based on the pool of available information.

EXAMPLE Consider that one's knowledge of how input estimate x_i was determined and how its standard uncertainty $u(x_i)$ was evaluated leads one to judge that the value of $u(x_i)$ is reliable to about 25 percent. This may be taken to mean that the relative uncertainty is $\cdot\, u(x_i)/u(x_i) = 0{,}25$, and thus from Equation (G.3), $\nu_i = (0{,}25)^{-2}/2 = 8$. If instead one had judged the value of $u(x_i)$ to be reliable to only about 50 percent, then $\nu_i = 2$. (See also Table E.1 in Annex E.)

G.4.3 In the discussion in 4.3 and 4.4 of Type B evaluation of standard uncertainty from an *a priori* probability distribution, it was implicitly assumed that the value of $u(x_i)$ resulting from such an evaluation is exactly known. For example, when $u(x_i)$ is obtained from a rectangular probability distribution of assumed half-width $a = (a_+ - a_-)/2$ as in 4.3.7 and 4.4.5, $u(x_i) = a/\sqrt{3}$ is viewed as a constant with no uncertainty because a_+ and a_-, and thus a, are so viewed (but see 4.3.9, Note 2). This implies through Equation (G.3) that $\nu_i \to \infty$ or $1/\nu_i \to 0$, but it causes no difficulty in evaluating Equation (G.2b). Further, assuming that $\nu_i \to \infty$ is not necessarily unrealistic; it is common practice to choose a_- and a_+ in such a way that the probability of the quantity in question lying outside the interval a_- to a_+ is extremely small.

G.5 Other considerations

G.5.1 An expression found in the literature on measurement uncertainty and often used to obtain an uncertainty that is intended to provide an interval with a 95 percent level of confidence may be written as

$$U_{95} = \left[t_{95}^2(\nu_{\text{eff}})\, s^2 + 3u^2 \right]^{1/2} \tag{G.4}$$

Here $t_{95}(\nu_{\text{eff}})$ is taken from the t-distribution for ν_{eff} degrees of freedom and $p = 95$ percent; ν_{eff} is the effective degrees of freedom calculated from the Welch-Satterthwaite formula [Equation (G.2b)] taking into account *only* those standard uncertainty components s_i that have been evaluated statistically from repeated observations in the *current* measurement; $s^2 = \sum c_i^2 s_i^2$; $c_i = \partial f / \partial x_i$; and $u^2 = \sum u_j^2(y) = \sum c_j^2 \left(a_j^2 / 3 \right)$ accounts for *all* other components of uncertainty, where $+a_j$ and $-a_j$ are the assumed exactly known upper and lower bounds of X_j relative to its best estimate x_j (that is, $x_j - a_j \leqslant X_j \leqslant x_j + a_j$).

NOTE A component based on repeated observations made *outside* the current measurement is treated in the same way as any other component included in u^2. Hence, in order to make a meaningful comparison between Equation (G.4) and Equation (G.5) of the following subclause, it is assumed that such components, if present, are negligible.

G.5.2 If an expanded uncertainty that provides an interval with a 95 percent level of confidence is evaluated according to the methods recommended in G.3 and G.4, the resulting expression in place of Equation (G.4) is

$$U_{95} = t_{95}(\nu_{\text{eff}}) \left(s^2 + u^2 \right)^{1/2} \tag{G.5}$$

where ν_{eff} is calculated from Equation (G.2b) and the calculation includes *all* uncertainty components.

In most cases, the value of U_{95} from Equation (G.5) will be larger than the value of U_{95} from Equation (G.4), if it is assumed that in evaluating Equation (G.5), all Type B variances are obtained from *a priori* rectangular distributions with half-widths that are the same as the bounds a_j used to compute u^2 of Equation (G.4). This may be understood by recognizing that, although $t_{95}(\nu_{\text{eff}})$ will in most cases be somewhat larger than $t_{95}(\nu_{\text{eff}})$, both factors are close to 2; and in Equation (G.5) u^2 is multiplied by $t_p^2(\nu_{\text{eff}})$ 4 while in Equation (G.4) it is multiplied by 3. Although the two expressions yield equal values of U_{95} and U_{95} for $u^2 \ll s^2$, U_{95} will be as much as 13 percent smaller than U_{95} if $u^2 \gg s^2$. Thus in general, Equation (G.4) yields an uncertainty that provides an interval having a *smaller* level of confidence than the interval provided by the expanded uncertainty calculated from Equation (G.5).

NOTE 1 In the limits $u^2/s^2 \to 8$ and $\nu_{\text{eff}} \to 8$, $U_{95} \to 1{,}732u$ while $U_{95} \to 1{,}960u$. In this case, U_{95} provides an interval having only a 91,7 percent level of confidence, while U_{95} provides a 95 percent interval. This case is approximated in practice when the components obtained from estimates of upper and lower bounds are dominant, large in number, and have values of $u_j^2(y) = c_j^2 a_j^2 / 3$ that are of comparable size.

NOTE 2 For a normal distribution, the coverage factor $k = \sqrt{3}$ 1,732 provides an interval with a level of confidence p = 91,673... percent. This value of p is robust in the sense that it is, in comparison with that of any other value, optimally independent of small deviations of the input quantities from normality.

G.5.3 Occasionally an input quantity X_i is distributed asymmetrically — deviations about its expected value of one sign are more probable than deviations of the opposite sign (see 4.3.8). Although this makes no difference in the evaluation of the standard uncertainty $u(x_i)$ of the estimate x_i of X_i, and thus in the evaluation of $u_c(y)$, it may affect the calculation of U.

It is usually convenient to give a symmetric interval, $Y = y \pm U$, unless the interval is such that there is a cost differential between deviations of one sign over the other. If the asymmetry of X_i causes only a small asymmetry in the probability distribution characterized by the measurement result y and its combined standard uncertainty $u_c(y)$, the probability lost on one side by quoting a symmetric interval is compensated by the probability gained on the other side. The alternative is to give an interval that is symmetric in probability (and thus asymmetric in U): the probability that Y lies below the lower limit $y - U_-$ is equal to the probability that Y lies above the upper limit $y + U_+$. But in order to quote such limits, more information than simply the estimates y and $u_c(y)$ [and hence more information than simply the estimates x_i and $u(x_i)$ of each input quantity X_i] is needed.

G.5.4 The evaluation of the expanded uncertainty U_p given here in terms of $u_c(y)$, ν_{eff}, and the factor $t_p(\nu_{eff})$ from the t-distribution is only an approximation, and it has its limitations. The distribution of $(y - Y)/u_c(y)$ is given by the t-distribution only if the distribution of Y is normal, the estimate y and its combined standard uncertainty $u_c(y)$ are independent, and if the distribution of $u_c^2(y)$ is a . ^{2} distribution. The introduction of ν_{eff}, Equation (G.2b), deals only with the latter problem, and provides an approximately . ^{2} distribution for $u_c^2(y)$; the other part of the problem, arising from the non-normality of the distribution of Y, requires the consideration of higher moments in addition to the variance.

G.6 Summary and conclusions

G.6.1 The coverage factor k_p that provides an interval having a level of confidence p close to a specified level can only be found if there is extensive knowledge of the probability distribution of each input quantity and if these distributions are combined to obtain the distribution of the output quantity. The input estimates x_i and their standard uncertainties $u(x_i)$ by themselves are inadequate for this purpose.

G.6.2 Because the extensive computations required to combine probability distributions are seldom justified by the extent and reliability of the available information, an approximation to the distribution of the output quantity is acceptable. Because of the Central Limit Theorem, it is usually sufficient to assume that the probability distribution of $(y - Y)/u_c(y)$ is the t-distribution and take $k_p = t_p(\nu_{eff})$, with the t-factor based on an effective degrees of freedom ν_{eff} of $u_c(y)$ obtained from the Welch-Satterthwaite formula, Equation (G.2b).

G.6.3 To obtain ν_{eff} from Equation (G.2b) requires the degrees of freedom ν_i for each standard uncertainty component. For a component obtained from a Type A evaluation, ν_i is obtained from the number of independent repeated observations upon which the corresponding input estimate is based and the number of independent quantities determined from those observations (see G.3.3). For a component obtained from a Type B evaluation, ν_i is obtained from the judged reliability of the value of that component [see G.4.2 and Equation (G.3)].

G.6.4 Thus the following is a summary of the preferred method of calculating an expanded uncertainty $U_p = k_p u_c(y)$ intended to provide an interval $Y = y \pm U_p$ that has an approximate level of confidence p:

1) Obtain y and $u_c(y)$ as described in Clauses 4 and 5.

2) Compute ν_{eff} from the Welch-Satterthwaite formula, Equation (G.2b) (repeated here for easy reference)

$$\nu_{eff} = \frac{u_c^4(y)}{\sum_{i=1}^{N} \frac{u_i^4(y)}{\nu_i}} \qquad \text{(G.2b)}$$

If $u(x_i)$ is obtained from a Type A evaluation, determine ν_i as outlined in G.3.3. If $u(x_i)$ is obtained from a Type B evaluation and it can be treated as exactly known, which is often the case in practice, $\nu_i \to 8$; otherwise, estimate ν_i from Equation (G.3).

3) Obtain the t-factor $t_p(\nu_{eff})$ for the desired level of confidence p from Table G.2. If ν_{eff} is not an integer, either interpolate or truncate ν_{eff} to the next lower integer.

4) Take $k_p = t_p(\nu_{eff})$ and calculate $U_p = k_p u_c(y)$.

G.6.5 In certain situations, which should not occur too frequently in practice, the conditions required by the Central Limit Theorem may not be well met and the approach of G.6.4 may lead to an unacceptable result. For example, if $u_c(y)$ is dominated by a component of uncertainty evaluated from a rectangular distribution whose bounds are assumed to be exactly known, it is possible [if $t_p(\nu_{eff}) > \sqrt{3}$] that $y + U_p$ and $y - U_p$, the upper and lower limits of the interval defined by U_p, could lie outside the bounds of the probability distribution of the output quantity Y. Such cases must be dealt with on an individual basis but are often amenable to an approximate analytic treatment (involving, for example, the convolution of a normal distribution with a rectangular distribution [10]).

G.6.6 For many practical measurements in a broad range of fields, the following conditions prevail:

- the estimate y of the measurand Y is obtained from estimates x_i of a significant number of input quantities X_i that are describable by well-behaved probability distributions, such as the normal and rectangular distributions;

- the standard uncertainties $u(x_i)$ of these estimates, which may be obtained from either Type A or Type B evaluations, contribute comparable amounts to the combined standard uncertainty $u_c(y)$ of the measurement result y;

- the linear approximation implied by the law of propagation of uncertainty is adequate (see 5.1.2 and E.3.1);

- the uncertainty of $u_c(y)$ is reasonably small because its effective degrees of freedom ν_{eff} has a significant magnitude, say greater than 10.

Under these circumstances, the probability distribution characterized by the measurement result and its combined standard uncertainty can be assumed to be normal because of the Central Limit Theorem; and $u_c(y)$ can be taken as a reasonably reliable estimate of the standard deviation of that normal distribution because of the significant size of ν_{eff}. Then, based on the discussion given in this annex, including that emphasizing the approximate nature of the uncertainty evaluation process and the impracticality of trying to distinguish between intervals having levels of confidence that differ by one or two percent, one may do the following:

- adopt $k = 2$ and assume that $U = 2u_c(y)$ defines an interval having a level of confidence of approximately 95 percent;

or, for more critical applications,

- adopt $k = 3$ and assume that $U = 3u_c(y)$ defines an interval having a level of confidence of approximately 99 percent.

Although this approach should be suitable for many practical measurements, its applicability to any particular measurement will depend on how close $k = 2$ must be to $t_{95}(\nu_{eff})$ or $k = 3$ must be to $t_{99}(\nu_{eff})$; that is, on how close the level of confidence of the interval defined by $U = 2u_c(y)$ or $U = 3u_c(y)$ must be to 95 percent or 99 percent, respectively. Although for $\nu_{eff} = 11$, $k = 2$ and $k = 3$ underestimate $t_{95}(11)$ and $t_{99}(11)$ by only about 10 percent and 4 percent, respectively (see Table G.2), this may not be acceptable in some cases. Further, for all values of ν_{eff} somewhat larger than 13, $k = 3$ produces an interval having a level of confidence larger than 99 percent. (See Table G.2, which also shows that for $\nu_{eff} \to \infty$ the levels of confidence of the intervals produced by $k = 2$ and $k = 3$ are 95,45 percent and 99,73 percent, respectively). Thus, in practice, the size of ν_{eff} and what is required of the expanded uncertainty will determine whether this approach can be used.

Table G.2 — Value of $t_p(v)$ from the t-distribution for degrees of freedom v that defines an interval $-t_p(v)$ to $+t_p(v)$ that encompasses the fraction p of the distribution

Degrees of freedom	Fraction p in percent					
v	68,27 [a]	90	95	95,45 [a]	99	99,73 [a]
1	1,84	6,31	12,71	13,97	63,66	235,80
2	1,32	2,92	4,30	4,53	9,92	19,21
3	1,20	2,35	3,18	3,31	5,84	9,22
4	1,14	2,13	2,78	2,87	4,60	6,62
5	1,11	2,02	2,57	2,65	4,03	5,51
6	1,09	1,94	2,45	2,52	3,71	4,90
7	1,08	1,89	2,36	2,43	3,50	4,53
8	1,07	1,86	2,31	2,37	3,36	4,28
9	1,06	1,83	2,26	2,32	3,25	4,09
10	1,05	1,81	2,23	2,28	3,17	3,96
11	1,05	1,80	2,20	2,25	3,11	3,85
12	1,04	1,78	2,18	2,23	3,05	3,76
13	1,04	1,77	2,16	2,21	3,01	3,69
14	1,04	1,76	2,14	2,20	2,98	3,64
15	1,03	1,75	2,13	2,18	2,95	3,59
16	1,03	1,75	2,12	2,17	2,92	3,54
17	1,03	1,74	2,11	2,16	2,90	3,51
18	1,03	1,73	2,10	2,15	2,88	3,48
19	1,03	1,73	2,09	2,14	2,86	3,45
20	1,03	1,72	2,09	2,13	2,85	3,42
25	1,02	1,71	2,06	2,11	2,79	3,33
30	1,02	1,70	2,04	2,09	2,75	3,27
35	1,01	1,70	2,03	2,07	2,72	3,23
40	1,01	1,68	2,02	2,06	2,70	3,20
45	1,01	1,68	2,01	2,06	2,69	3,18
50	1,01	1,68	2,01	2,05	2,68	3,16
100	1,005	1,660	1,984	2,025	2,626	3,077
∞	1,000	1,645	1,960	2,000	2,576	3,000

a) For a quantity z described by a normal distribution with expectation α_z and standard deviation σ, the interval $\alpha_z \pm k\sigma$ encompasses $p = 68{,}27$ percent, 95,45 percent and 99,73 percent of the distribution for $k = 1$, 2 and 3, respectively.

Annex H

Examples

This annex gives six examples, H.1 to H.6, which are worked out in considerable detail in order to illustrate the basic principles presented in this *Guide* for evaluating and expressing uncertainty in measurement. Together with the examples included in the main text and in some of the other annexes, they should enable the users of this *Guide* to put these principles into practice in their own work.

Because the examples are for illustrative purposes, they have by necessity been simplified. Moreover, because they and the numerical data used in them have been chosen mainly to demonstrate the principles of this *Guide*, neither they nor the data should necessarily be interpreted as describing real measurements. While the data are used as given, in order to prevent rounding errors, more digits are retained in intermediate calculations than are usually shown. Thus the stated result of a calculation involving several quantities may differ slightly from the result implied by the numerical values given in the text for these quantities.

It is pointed out in earlier portions of this *Guide* that classifying the methods used to evaluate components of uncertainty as Type A or Type B is for convenience only; it is not required for the determination of the combined standard uncertainty or expanded uncertainty of a measurement result because all uncertainty components, however they are evaluated, are treated in the same way (see 3.3.4, 5.1.2, and E.3.7). Thus, in the examples, the method used to evaluate a particular component of uncertainty is not specifically identified as to its type. However, it will be clear from the discussion whether a component is obtained from a Type A or a Type B evaluation.

H.1 End-gauge calibration

This example demonstrates that even an apparently simple measurement may involve subtle aspects of uncertainty evaluation.

H.1.1 The measurement problem

The length of a nominally 50 mm end gauge is determined by comparing it with a known standard of the same nominal length. The direct output of the comparison of the two end gauges is the difference d in their lengths:

$$d = l\left(1+\alpha_{\cdot}\ \right) - l_{S}\left(1+\alpha_{S\cdot\ S}\right) \tag{H.1}$$

where

l is the measurand, that is, the length at 20 °C of the end gauge being calibrated;

l_S is the length of the standard at 20 °C as given in its calibration certificate;

α and α_S are the coefficients of thermal expansion, respectively, of the gauge being calibrated and the standard;

$_{\cdot}$ and $_{\cdot\ S}$ are the *deviations* in temperature from the 20 °C reference temperature, respectively, of the gauge and the standard.

H.1.2 Mathematical model

From Equation (H.1), the measurand is given by

$$l = \frac{l_S\left(1+\alpha_{S\cdot}{}_S\right)+d}{\left(1+\alpha_{\cdot}{}_{\cdot}\right)} = l_S + d + l_S\left(\alpha_{S\cdot}{}_S - \alpha_{\cdot}{}_{\cdot}\right) + \ldots \tag{H.2}$$

If the difference in temperature between the end gauge being calibrated and the standard is written as $\delta_{\cdot} = {}_{\cdot} - {}_{\cdot}{}_S$, and the difference in their thermal expansion coefficients as $\delta\alpha = \alpha - \alpha_S$, Equation (H.2) becomes

$$l = f\left(l_S, d, \alpha_S, {}_{\cdot}{}_{\cdot}, \delta\alpha, \delta_{\cdot}{}_{\cdot}\right) = l_S + d - l_S\left(\delta\alpha \bullet_{\cdot}{}_{\cdot} + \alpha_S \bullet \delta_{\cdot}{}_{\cdot}\right) \tag{H.3}$$

The differences $\delta_{\cdot}{}_{\cdot}$ and $\delta\alpha$, but not their uncertainties, are estimated to be zero; and $\delta\alpha$, α_S, $\delta_{\cdot}{}_{\cdot}$, and ${}_{\cdot}{}_{\cdot}$ are assumed to be uncorrelated. (If the measurand were expressed in terms of the variables ${}_{\cdot}{}_{\cdot}$, ${}_{\cdot}{}_S$, α, and α_S, it would be necessary to include the correlation between ${}_{\cdot}{}_{\cdot}$ and ${}_{\cdot}{}_S$, and between α and α_S.)

It thus follows from Equation (H.3) that the estimate of the value of the measurand l may be obtained from the simple expression $l_S + \bar{d}$, where l_S is the length of the standard at 20 °C as given in its calibration certificate and d is estimated by $\bar{d}$, the arithmetic mean of $n = 5$ independent repeated observations. The combined standard uncertainty $u_c(l)$ of l is obtained by applying Equation (10) in 5.1.2 to Equation (H.3), as discussed below.

NOTE In this and the other examples, for simplicity of notation, the same symbol is used for a quantity and its estimate.

H.1.3 Contributory variances

The pertinent aspects of this example as discussed in this and the following subclauses are summarized in Table H.1.

Since it is assumed that $\delta\alpha = 0$ and $\delta_{\cdot}{}_{\cdot} = 0$, the application of Equation (10) in 5.1.2 to Equation (H.3) yields

$$u_c^2(l) = c_S^2 u^2(l_S) + c_d^2 u^2(d) + c_{\alpha_S}^2 u^2(\alpha_S) + c_{\cdot}^2 u^2({}_{\cdot}{}_{\cdot}) + c_{\delta\alpha}^2 u^2(\delta\alpha) + c_{\delta_{\cdot}{}_{\cdot}}^2 u^2(\delta_{\cdot}{}_{\cdot}) \tag{H.4}$$

with

$$c_S = \partial f / \partial l_S = 1 - \left(\delta\alpha \bullet_{\cdot}{}_{\cdot} + \alpha_S \bullet \delta_{\cdot}{}_{\cdot}\right) = 1$$

$$c_d = \partial f / \partial d = 1$$

$$c_{\alpha_S} = \partial f / \partial \alpha_S = -l_S \delta_{\cdot}{}_{\cdot} = 0$$

$$c_{\cdot} = \partial f / \partial {}_{\cdot}{}_{\cdot} = -l_S \delta\alpha = 0$$

$$c_{\delta\alpha} = \partial f / \partial \delta\alpha = -l_S{}_{\cdot}$$

$$c_{\delta_{\cdot}{}_{\cdot}} = \partial f / \partial \delta_{\cdot}{}_{\cdot} = -l_S \alpha_S$$

and thus

$$u_c^2(l) = u^2(l_S) + u^2(d) + l_S^2{}_{\cdot}{}^2 u^2(\delta\alpha) + l_S^2 \alpha_S^2 u^2(\delta_{\cdot}{}_{\cdot}) \tag{H.5}$$

H.1.3.1 Uncertainty of the calibration of the standard, $u(l_S)$

The calibration certificate gives as the expanded uncertainty of the standard $U = 0{,}075$ µm and states that it was obtained using a coverage factor of $k = 3$. The standard uncertainty is then

$$u(l_S) = (0{,}075\ \mu\text{m})/3 = 25\ \text{nm}$$

H.1.3.2 Uncertainty of the measured difference in lengths, $u(d)$

The pooled experimental standard deviation characterizing the comparison of l and l_S was determined from the variability of 25 independent repeated observations of the difference in lengths of two standard end gauges and was found to be 13 nm. In the comparison of this example, five repeated observations were taken. The standard uncertainty associated with the arithmetic mean of these readings is then (see 4.2.4)

$$u(\bar{d}) = s(\bar{d}) = (13\ \text{nm})/\sqrt{5} = 5{,}8\ \text{nm}$$

According to the calibration certificate of the comparator used to compare l with l_S, its uncertainty "due to random errors" is $\pm 0{,}01$ µm at a level of confidence of 95 percent and is based on 6 replicate measurements; thus the standard uncertainty, using the t-factor $t_{95}(5) = 2{,}57$ for $\nu = 6 - 1 = 5$ degrees of freedom (see Annex G, Table G.2), is

$$u(d_1) = (0{,}01\ \mu\text{m})/2{,}57 = 3{,}9\ \text{nm}$$

The uncertainty of the comparator "due to systematic errors" is given in the certificate as 0,02 µm at the "three sigma level". The standard uncertainty from this cause may therefore be taken to be

$$u(d_2) = (0{,}02\ \mu\text{m})/3 = 6{,}7\ \text{nm}$$

The total contribution is obtained from the sum of the estimated variances:

$$u^2(d) = u^2(\bar{d}) + u^2(d_1) + u^2(d_2) = 93\ \text{nm}^2$$

or

$$u(d) = 9{,}7\ \text{nm}$$

H.1.3.3 Uncertainty of the thermal expansion coefficient, $u(\alpha_S)$

The coefficient of thermal expansion of the standard end gauge is given as $\alpha_S = 11{,}5 \cdot 10^{-6}\ °\text{C}^{-1}$ with an uncertainty represented by a rectangular distribution with bounds $\pm 2 \cdot 10^{-6}\ °\text{C}^{-1}$. The standard uncertainty is then [see Equation (7) in 4.3.7]

$$u(\alpha_S) = \left(2 \cdot 10^{-6}\ °\text{C}^{-1}\right)/\sqrt{3} = 1{,}2 \cdot 10^{-6}\ °\text{C}^{-1}$$

Since $c_{\alpha_S} = \partial f / \partial \alpha_S = -l_S \delta_{\cdot} = 0$ as indicated in H.1.3, this uncertainty contributes nothing to the uncertainty of l in first order. It does, however, have a second-order contribution that is discussed in H.1.7.

Table H.1 — Summary of standard uncertainty components

Standard uncertainty component $u(x_i)$	Source of uncertainty	Value of standard uncertainty $u(x_i)$	$c_i = \partial f / \partial x_i$	$u_i(l) = \lvert c_i \rvert u(x_i)$ (nm)	Degrees of freedom
$u(l_S)$	Calibration of standard end gauge	25 nm	1	25	18
$u(d)$	Measured difference between end gauges	9,7 nm	1	9,7	25,6
$u(\bar{d})$	repeated observations	5,8 nm			24
$u(d_1)$	random effects of comparator	3,9 nm			5
$u(d_2)$	systematic effects of comparator	6,7 nm			8
$u(\alpha_S)$	Thermal expansion coefficient of standard end gauge	$1{,}2 \cdot 10^{-6}$ °C^{-1}	0	0	
$u(.\ )$	Temperature of test bed	0,41 °C	0	0	
$u(\bar{.})$	mean temperature of bed	0,2 °C			
$u(.\ )$	cyclic variation of temperature of room	0,35 °C			
$u(\delta\alpha)$	Difference in expansion coefficients of end gauges	$0{,}58 \cdot 10^{-6}$ °C^{-1}	$-l_S$.	2,9	50
$u(\delta.\ )$	Difference in temperatures of end gauges	0,029 °C	$-l_S\alpha_S$	16,6	2
				$u_c^2(l) = \sum u_i^2(l) = 1\,002$ nm^2	
				$u_c(l) = 32$ nm	
				$v_{eff}(l) = 16$	

H.1.3.4 Uncertainty of the deviation of the temperature of the end gauge, $u(.\ )$

The temperature of the test bed is reported as $(19{,}9 \pm 0{,}5)$ °C; the temperature at the time of the individual observations was not recorded. The stated maximum offset, $. = 0{,}5$ °C, is said to represent the amplitude of an approximately cyclical variation of the temperature under a thermostatic system, not the uncertainty of the mean temperature. The value of the mean temperature deviation

$$\bar{.} = 19{,}9\ °C - 20\ °C = -0{,}1\ °C$$

is reported as having a standard uncertainty itself due to the uncertainty in the mean temperature of the test bed of

$$u(\bar{.}) = 0{,}2\ °C$$

while the cyclic variation in time produces a U-shaped (arcsine) distribution of temperatures resulting in a standard uncertainty of

$$u(.\ ) = (0{,}5\ °C)/\sqrt{2} = 0{,}35\ °C$$

The temperature deviation $\delta\theta$ may be taken equal to $\delta\bar{\theta}$, and the standard uncertainty of θ is obtained from

$$u^2(\delta\theta) = u^2(\delta\bar{\theta}) + u^2(\delta\theta) = 0{,}165\ °C^2$$

which gives

$$u(\delta\theta) = 0{,}41\ °C$$

Since $c_\theta = \partial f / \partial\delta\theta = -l_S\delta\alpha = 0$ as indicated in H.1.3, this uncertainty also contributes nothing to the uncertainty of l in first order; but it does have a second-order contribution that is discussed in H.1.7.

H.1.3.5 Uncertainty of the difference in expansion coefficients, $u(\delta\alpha)$

The estimated bounds on the variability of $\delta\alpha$ are $\pm 1 \cdot 10^{-6}\ °C^{-1}$, with an equal probability of $\delta\alpha$ having any value within those bounds. The standard uncertainty is

$$u(\delta\alpha) = \left(1 \cdot 10^{-6}\ °C^{-1}\right) / \sqrt{3} = 0{,}58 \cdot 10^{-6}\ °C^{-1}$$

H.1.3.6 Uncertainty of the difference in temperature of the gauges, $u(\delta\theta)$

The standard and the test gauge are expected to be at the same temperature, but the temperature difference could lie with equal probability anywhere in the estimated interval –0,05 °C to +0,05 °C. The standard uncertainty is

$$u(\delta\theta) = (0{,}05\ °C) / \sqrt{3} = 0{,}029\ °C$$

H.1.4 Combined standard uncertainty

The combined standard uncertainty $u_c(l)$ is calculated from Equation (H.5). The individual terms are collected and substituted into this expression to obtain

$$u_c^2(l) = (25\ \text{nm})^2 + (9{,}7\ \text{nm})^2 + (0{,}05\ \text{m})^2 (-0{,}1\ °C)^2 \left(0{,}58 \cdot 10^{-6}\ °C^{-1}\right)^2 + (0{,}05\ \text{m})^2 \left(11{,}5 \cdot 10^{-6}\ °C^{-1}\right)^2 (0{,}029\ °C)^2 \quad \text{(H.6a)}$$

$$= (25\ \text{nm})^2 + (9{,}7\ \text{nm})^2 + (2{,}9\ \text{nm})^2 + (16{,}6\ \text{nm})^2 = 1\,002\ \text{nm}^2 \quad \text{(H.6b)}$$

or

$$u_c(l) = 32\ \text{nm} \quad \text{(H.6c)}$$

The dominant component of uncertainty is obviously that of the standard, $u(l_S)$ = 25 nm.

H.1.5 Final result

The calibration certificate for the standard end gauge gives l_S = 50,000 623 mm as its length at 20 °C. The arithmetic mean $\bar{d}$ of the five repeated observations of the difference in lengths between the unknown end gauge and the standard gauge is 215 nm. Thus, since $l = l_S + \bar{d}$ (see H.1.2), the length l of the unknown end gauge at 20 °C is 50,000 838 mm. Following 7.2.2, the final result of the measurement may be stated as:

> l = 50,000 838 mm with a combined standard uncertainty u_c = 32 nm. The corresponding relative combined standard uncertainty is $u_c/l = 6{,}4 \cdot 10^{-7}$.

H.1.6 Expanded uncertainty

Suppose that one is required to obtain an expanded uncertainty $U_{99} = k_{99}u_c(l)$ that provides an interval having a level of confidence of approximately 99 percent. The procedure to use is that summarized in G.6.4, and the required degrees of freedom are indicated in Table H.1. These were obtained as follows:

1) *Uncertainty of the calibration of the standard*, $u(l_S)$ [H.1.3.1]. The calibration certificate states that the effective degrees of freedom of the combined standard uncertainty from which the quoted expanded uncertainty was obtained is $\nu_{\text{eff}}(l_S) = 18$.

2) *Uncertainty of the measured difference in lengths*, $u(d)$ [H.1.3.2]. Although $\bar{d}$ was obtained from five repeated observations, because $u(\bar{d})$ was obtained from a pooled experimental standard deviation based on 25 observations, the degrees of freedom of $u(\bar{d})$ is $\nu(\bar{d}) = 25 - 1 = 24$ (see H.3.6, note). The degrees of freedom of $u(d_1)$, the uncertainty due to random effects on the comparator, is $\nu(d_1) = 6 - 1 = 5$ because d_1 was obtained from six repeated measurements. The ±0,02 µm uncertainty for systematic effects on the comparator may be assumed to be reliable to 25 percent, and thus the degrees of freedom from Equation (G.3) in G.4.2 is $\nu(d_2) = 8$ (see the example of G.4.2). The effective degrees of freedom of $u(d)$, $\nu_{\text{eff}}(d)$, is then obtained from Equation (G.2b) in G.4.1:

$$\nu_{\text{eff}}(d) = \frac{\left[u^2(\bar{d}) + u^2(d_1) + u^2(d_2)\right]^2}{\dfrac{u^4(\bar{d})}{\nu(\bar{d})} + \dfrac{u^4(d_1)}{\nu(d_1)} + \dfrac{u^4(d_2)}{\nu(d_2)}} = \frac{(9{,}7\ \text{nm})^4}{\dfrac{(5{,}8\ \text{nm})^4}{24} + \dfrac{(3{,}9\ \text{nm})^4}{5} + \dfrac{(6{,}7\ \text{nm})^4}{8}} = 25{,}6$$

3) *Uncertainty of the difference in expansion coeficients*, $u(\delta\alpha)$ [H.1.3.5]. The estimated bounds of $\pm 1 \cdot 10^{-6}\ °\text{C}^{-1}$ on the variability of $\delta\alpha$ are deemed to be reliable to 10 percent. This gives, from Equation (G.3) in G.4.2, $\nu(\delta\alpha) = 50$.

4) *Uncertainty of the difference in temperatures of the gauges*, $u(\delta.\,)$ [H.1.3.6]. The estimated interval −0,05 °C to +0,05 °C for the temperature difference $\delta.$ is believed to be reliable only to 50 percent, which from Equation (G.3) in G.4.2 gives $\nu(\delta.\,) = 2$.

The calculation of $\nu_{\text{eff}}(l)$ from Equation (G.2b) in G.4.1 proceeds in exactly the same way as for the calculation of $\nu_{\text{eff}}(d)$ in 2) above. Thus from Equations (H.6b) and (H.6c) and the values for ν given in 1) through 4),

$$\nu_{\text{eff}}(l) = \frac{(32\ \text{nm})^4}{\dfrac{(25\ \text{nm})^4}{18} + \dfrac{(9{,}7\ \text{nm})^4}{25{,}6} + \dfrac{(2{,}9\ \text{nm})^4}{50} + \dfrac{(16{,}6\ \text{nm})^4}{2}} = 16{,}7$$

To obtain the required expanded uncertainty, this value is first truncated to the next lower integer, $\nu_{\text{eff}}(l) = 16$. It then follows from Table G.2 in Annex G that $t_{99}(16) = 2{,}92$, and hence $U_{99} = t_{99}(16)u_c(l) = 2{,}92 \cdot (32\ \text{nm}) = 93\ \text{nm}$. Following 7.2.4, the final result of the measurement may be stated as:

> l = (50,000 838 ± 0,000 093) mm, where the number following the symbol ± is the numerical value of an expanded uncertainty $U = ku_c$, with U determined from a combined standard uncertainty $u_c = 32$ nm and a coverage factor $k = 2{,}92$ based on the t-distribution for $\nu = 16$ degrees of freedom, and defines an interval estimated to have a level of confidence of 99 percent. The corresponding relative expanded uncertainty is $U/l = 1{,}9 \cdot 10^{-6}$.

H.1.7 Second-order terms

The note to 5.1.2 points out that Equation (10), which is used in this example to obtain the combined standard uncertainty $u_c(l)$, must be augmented when the nonlinearity of the function $Y = f(X_1, X_2, ..., X_N)$ is so significant that the higher-order terms in the Taylor series expansion cannot be neglected. Such is the case in this example, and therefore the evaluation of $u_c(l)$ as presented up to this point is not complete. Application to Equation (H.3) of the expression given in the note to 5.1.2 yields in fact two distinct non-negligible second-order terms to be added to Equation (H.5). These terms, which arise from the quadratic term in the expression of the note, are

$$l_S^2 u^2(\delta\alpha) u^2(\,.\,) + l_S^2 u^2(\alpha_S) u^2(\delta.\,)$$

but only the first of these terms contributes significantly to $u_c(l)$:

$$l_S u(\delta\alpha) u(\,.\,) = (0{,}05\ \text{m})\left(0{,}58 \cdot 10^{-6}\ ^\circ\text{C}^{-1}\right)(0{,}41\ ^\circ\text{C}) = 11{,}7\ \text{nm}$$

$$l_S u(\alpha_S) u(\delta.\,) = (0{,}05\ \text{m})\left(1{,}2 \cdot 10^{-6}\ ^\circ\text{C}^{-1}\right)(0{,}029\ ^\circ\text{C}) = 1{,}7\ \text{nm}$$

The second-order terms increase $u_c(l)$ from 32 nm to 34 nm.

H.2 Simultaneous resistance and reactance measurement

This example demonstrates the treatment of multiple measurands or output quantities determined simultaneously in the same measurement and the correlation of their estimates. It considers only the random variations of the observations; in actual practice, the uncertainties of corrections for systematic effects would also contribute to the uncertainty of the measurement results. The data are analysed in two different ways with each yielding essentially the same numerical values.

H.2.1 The measurement problem

The resistance R and the reactance X of a circuit element are determined by measuring the amplitude V of a sinusoidally-alternating potential difference across its terminals, the amplitude I of the alternating current passing through it, and the phase-shift angle ϕ of the alternating potential difference relative to the alternating current. Thus the three input quantities are V, I, and ϕ and the three output quantities — the measurands — are the three impedance components R, X, and Z. Since $Z^2 = R^2 + X^2$, there are only two independent output quantities.

H.2.2 Mathematical model and data

The measurands are related to the input quantities by Ohm's law:

$$R = \frac{V}{I}\cos\phi; \quad X = \frac{V}{I}\sin\phi; \quad Z = \frac{V}{I} \tag{H.7}$$

Consider that five independent sets of simultaneous observations of the three input quantities V, I, and ϕ are obtained under similar conditions (see B.2.15), resulting in the data given in Table H.2. The arithmetic means of the observations and the experimental standard deviations of those means calculated from Equations (3) and (5) in 4.2 are also given. The means are taken as the best estimates of the expected values of the input quantities, and the experimental standard deviations are the standard uncertainties of those means.

Because the means $\bar{V}$, $\bar{I}$, and $\bar{\phi}$ are obtained from simultaneous observations, they are correlated and the correlations must be taken into account in the evaluation of the standard uncertainties of the measurands R, X, and Z. The required correlation coefficients are readily obtained from Equation (14) in 5.2.2 using values of $s(\bar{V}, \bar{I})$, $s(\bar{V}, \bar{\phi})$, and $s(\bar{I}, \bar{\phi})$ calculated from Equation (17) in 5.2.3. The results are included in Table H.2, where it should be recalled that $r(x_i, x_j) = r(x_j, x_i)$ and $r(x_i, x_i) = 1$.

Table H.2 — Values of the input quantities V, I, and ϕ obtained from five sets of simultaneous observations

Set number k	Input quantities		
	V (V)	I (mA)	ϕ (rad)
1	5,007	19,663	1,045 6
2	4,994	19,639	1,043 8
3	5,005	19,640	1,046 8
4	4,990	19,685	1,042 8
5	4,999	19,678	1,043 3
Arithmetic mean	$\bar{V} = 4{,}999\ 0$	$\bar{I} = 19{,}661\ 0$	$\bar{\phi} = 1{,}044\ 46$
Experimental standard deviation of mean	$s(\bar{V}) = 0{,}003\ 2$	$s(\bar{I}) = 0{,}009\ 5$	$s(\bar{\phi}) = 0{,}000\ 75$
Correlation coefficients			
$r(\bar{V}, \bar{I}) = -0{,}36$ $r(\bar{V}, \bar{\phi}) = 0{,}86$ $r(\bar{I}, \bar{\phi}) = -0{,}65$			

H.2.3 Results: approach 1

Approach 1 is summarized in Table H.3.

The values of the three measurands R, X, and Z are obtained from the relations given in Equation (H.7) using the mean values $\bar{V}$, $\bar{I}$, and $\bar{\phi}$ of Table H.2 for V, I, and ϕ. The standard uncertainties of R, X, and Z are obtained from Equation (16) in 5.2.2 since, as pointed out above, the input quantities $\bar{V}$, $\bar{I}$, and $\bar{\phi}$ are correlated. As an example, consider $Z = \bar{V}/\bar{I}$. Identifying $\bar{V}$ with x_1, $\bar{I}$ with x_2, and f with $Z = \bar{V}/\bar{I}$, Equation (16) in 5.2.2 yields for the combined standard uncertainty of Z

$$u_c^2(Z) = \left(\frac{1}{\bar{I}}\right)^2 u^2(\bar{V}) + \left(\frac{\bar{V}}{\bar{I}^2}\right)^2 u^2(\bar{I}) + 2\left(\frac{1}{\bar{I}}\right)\left(-\frac{\bar{V}}{\bar{I}^2}\right) u(\bar{V})\,u(\bar{I})\,r(\bar{V}, \bar{I}) \tag{H.8a}$$

$$= Z^2\left[\frac{u(\bar{V})}{\bar{V}}\right]^2 + Z^2\left[\frac{u(\bar{I})}{\bar{I}}\right]^2 - 2Z^2\left[\frac{u(\bar{V})}{\bar{V}}\right]\left[\frac{u(\bar{I})}{\bar{I}}\right] r(\bar{V}, \bar{I}) \tag{H.8b}$$

or

$$u_{c,r}^2(\bar{Z}) = u_r^2(\bar{V}) + u_r^2(\bar{I}) - 2u_r(\bar{V})\,u_r(\bar{I})\,r(\bar{V}, \bar{I}) \tag{H.8c}$$

where $u(\bar{V}) = s(\bar{V})$, $u(\bar{I}) = s(\bar{I})$, and the subscript "r" in the last expression indicates that u is a relative uncertainty. Substitution of the appropriate values from Table H.2 into Equation (H.8a) then gives $u_c(Z) = 0{,}236$ . .

Because the three measurands or output quantities depend on the same input quantities, they too are correlated. The elements of the covariance matrix that describes this correlation may be written in general as

$$u(y_l, y_m) = \sum_{i=1}^{N}\sum_{j=1}^{N} \frac{\partial y_l}{\partial x_i}\frac{\partial y_m}{\partial x_j}\, u(x_i)\,u(x_j)\,r(x_i, x_j) \tag{H.9}$$

where $y_l = f_l(x_1, x_2, ..., x_N)$ and $y_m = f_m(x_1, x_2, ..., x_N)$. Equation (H.9) is a generalization of Equation (F.2) in F.1.2.3 when the q_l in that expression are correlated. The estimated correlation coefficients of the output quantities are given by $r(y_l, y_m) = u(y_l, y_m)/u(y_l)u(y_m)$, as indicated in Equation (14) in 5.2.2. It should be recognized that the diagonal elements of the covariance matrix, $u(y_l, y_l) = u^2(y_l)$, are the estimated variances of the output quantities y_l (see 5.2.2, Note 2) and that for $m = l$, Equation (H.9) is identical to Equation (16) in 5.2.2.

To apply Equation (H.9) to this example, the following identifications are made:

$$y_1 = R \qquad x_1 = \bar{V} \qquad u(x_i) = s(x_i)$$
$$y_2 = X \qquad x_2 = \bar{I} \qquad N = 3$$
$$y_3 = Z \qquad x_3 = \bar{\phi}$$

The results of the calculations of R, X, and Z and of their estimated variances and correlation coefficients are given in Table H.3.

Table H.3 — Calculated values of the output quantities R, X, and Z: approach 1

Measurand index l	Relationship between estimate of measurand y_l and input estimates x_i	Value of estimate y_l, which is the result of measurement	Combined standard uncertainty $u_c(y_l)$ of result of measurement
1	$y_1 = R = (\bar{V}/\bar{I})\cos\bar{\phi}$	$y_1 = R = 127{,}732\ \Omega$	$u_c(R) = 0{,}071\ \Omega$ $u_c(R)/R = 0{,}06 \cdot 10^{-2}$
2	$y_2 = X = (\bar{V}/\bar{I})\sin\bar{\phi}$	$y_2 = X = 219{,}847\ \Omega$	$u_c(X) = 0{,}295\ \Omega$ $u_c(X)/X = 0{,}13 \cdot 10^{-2}$
3	$y_3 = Z = \bar{V}/\bar{I}$	$y_3 = Z = 254{,}260\ \Omega$	$u_c(Z) = 0{,}236\ \Omega$ $u_c(Z)/Z = 0{,}09 \cdot 10^{-2}$
Correlation coefficients $r(y_l, y_m)$			
$r(y_1, y_2) = r(R, X) = -0{,}588$ $r(y_1, y_3) = r(R, Z) = -0{,}485$ $r(y_2, y_3) = r(X, Z) = 0{,}993$			

H.2.4 Results: approach 2

Approach 2 is summarized in Table H.4.

Since the data have been obtained as five sets of observations of the three input quantities V, I, and ϕ, it is possible to compute a value for R, X, and Z from *each set* of input data, and then take the arithmetic mean of the five individual values to obtain the best estimates of R, X, and Z. The experimental standard deviation of each mean (which is its combined standard uncertainty) is then calculated from the five individual values in the usual way [Equation (5) in 4.2.3]; and the estimated covariances of the three means are calculated by applying Equation (17) in 5.2.3 directly to the five individual values from which each mean is obtained. There are no differences in the output values, standard uncertainties, and estimated covariances provided by the two approaches except for second-order effects associated with replacing terms such as $\bar{V}/\bar{I}$ and $\cos\bar{\phi}$ by $\overline{V/I}$ and $\overline{\cos\phi}$.

To demonstrate this approach, Table H.4 gives the values of R, X and Z calculated from each of the five sets of observations. The arithmetic means, standard uncertainties, and estimated correlation coefficients are then directly computed from these individual values. The numerical results obtained in this way are negligibly different from the results given in Table H.3.

Table H.4 — Calculated values of the output quantities R, X, and Z: approach 2

Set number k	Individual values of measurands		
	$R = (V/I)\cos\phi$ (.)	$X = (V/I)\sin\phi$ (.)	$Z = V/I$ (.)
1	127,67	220,32	254,64
2	127,89	219,79	254,29
3	127,51	220,64	254,84
4	127,71	218,97	253,49
5	127,88	219,51	254,04
Arithmetic mean	$y_1 = \bar{R} = 127{,}732$	$y_2 = \bar{X} = 219{,}847$	$y_3 = \bar{Z} = 254{,}260$
Experimental standard deviation of mean	$s(\bar{R}) = 0{,}071$	$s(\bar{X}) = 0{,}295$	$s(\bar{Z}) = 0{,}236$
Correlation coefficients $r(y_l, y_m)$			
$r(y_1, y_2) = r(\bar{R}, \bar{X}) = -0{,}588$ $r(y_1, y_3) = r(\bar{R}, \bar{Z}) = -0{,}485$ $r(y_2, y_3) = r(\bar{X}, \bar{Z}) = 0{,}993$			

In the terminology of the Note to 4.1.4, approach 2 is an example of obtaining the estimate y from $\bar{Y} = \left(\sum_{k=1}^{n} Y_k\right)/n$, while approach 1 is an example of obtaining y from $y = f(\bar{X}_1, \bar{X}_2, ..., \bar{X}_N)$. As pointed out in that note, in general, the two approaches will give *identical* results if f is a linear function of its input quantities (provided that the experimentally observed correlation coefficients are taken into account when implementing approach 1). If f is not a linear function, then the results of approach 1 will differ from those of approach 2 depending on the degree of nonlinearity and the estimated variances and covariances of the X_i. This may be seen from the expression

$$y = f\left(\bar{X}_1, \bar{X}_2, ..., \bar{X}_N\right) + \frac{1}{2}\sum_{i=1}^{N}\sum_{j=1}^{N} \frac{\partial^2 f}{\partial \bar{X}_i \partial \bar{X}_j} u\left(\bar{X}_i, \bar{X}_j\right) + ... \qquad \text{(H.10)}$$

where the second term on the right-hand side is the second-order term in the Taylor series expansion of f in terms of the $\bar{X}_i$ (see also 5.1.2, note). In the present case, approach 2 is preferred because it avoids the approximation $y = f(\bar{X}_1, \bar{X}_2, ..., \bar{X}_N)$ and better reflects the measurement procedure used — the data were in fact collected in sets.

On the other hand, approach 2 would be inappropriate if the data of Table H.2 represented $n_1 = 5$ observations of the potential difference V, followed by $n_2 = 5$ observations of the current I, and then followed by $n_3 = 5$ observations of the phase ϕ, and would be impossible if $n_1 \neq n_2 \neq n_3$. (It is in fact poor measurement procedure to carry out the measurements in this way since the potential difference across a fixed impedance and the current through it are directly related.)

If the data of Table H.2 are reinterpreted in this manner so that approach 2 is inappropriate, and if correlations among the quantities V, I, and ϕ are assumed to be absent, then the observed correlation coefficients have no significance and should be set equal to zero. If this is done in Table H.2, Equation (H.9) reduces to the equivalent of Equation (F.2) in F.1.2.3, namely,

$$u\left(y_l, y_m\right) = \sum_{i=1}^{N} \frac{\partial y_l}{\partial x_i} \frac{\partial y_m}{\partial x_i} u^2\left(x_i\right) \qquad \text{(H.11)}$$

and its application to the data of Table H.2 leads to the changes in Table H.3 shown in Table H.5.

Table H.5 — Changes in Table H.3 under the assumption that the correlation coefficients of Table H.2 are zero

Combined standard uncertainty $u_c(y_l)$ of result of measurement
$u_c(R) = 0{,}195\ .$ $u_c(R)/R = 0{,}15 \cdot 10^{-2}$
$u_c(X) = 0{,}201\ .$ $u_c(X)/X = 0{,}09 \cdot 10^{-2}$
$u_c(Z) = 0{,}204\ .$ $u_c(Z)/Z = 0{,}08 \cdot 10^{-2}$
Correlation coefficients $r(y_l, y_m)$
$r(y_1, y_2) = r(R, X) = 0{,}056$ $r(y_1, y_3) = r(R, Z) = 0{,}527$ $r(y_2, y_3) = r(X, Z) = 0{,}878$

H.3 Calibration of a thermometer

This example illustrates the use of the method of least squares to obtain a linear calibration curve and how the parameters of the fit, the intercept and slope, and their estimated variances and covariance, are used to obtain from the curve the value and standard uncertainty of a predicted correction.

H.3.1 The measurement problem

A thermometer is calibrated by comparing $n = 11$ temperature readings t_k of the thermometer, each having negligible uncertainty, with corresponding known reference temperatures $t_{R,\,k}$ in the temperature range 21 °C to 27 °C to obtain the corrections $b_k = t_{R,\,k} - t_k$ to the readings. The *measured* corrections b_k and *measured* temperatures t_k are the input quantities of the evaluation. A linear calibration curve

$$b(t) = y_1 + y_2 (t - t_0) \tag{H.12}$$

is fitted to the measured corrections and temperatures by the method of least squares. The parameters y_1 and y_2, which are respectively the intercept and slope of the calibration curve, are the two measurands or output quantities to be determined. The temperature t_0 is a conveniently chosen exact reference temperature; it is not an independent parameter to be determined by the least-squares fit. Once y_1 and y_2 are found, along with their estimated variances and covariance, Equation (H.12) can be used to predict the value and standard uncertainty of the correction to be applied to the thermometer for any value t of the temperature.

H.3.2 Least-squares fitting

Based on the method of least squares and under the assumptions made in H.3.1 above, the output quantities y_1 and y_2 and their estimated variances and covariance are obtained by minimizing the sum

$$S = \sum_{k=1}^{n} \left[b_k - y_1 - y_2 (t_k - t_0) \right]^2$$

This leads to the following equations for y_1, y_2, their experimental variances $s^2(y_1)$ and $s^2(y_2)$, and their estimated correlation coefficient $r(y_1, y_2) = s(y_1, y_2)/s(y_1)s(y_2)$, where $s(y_1, y_2)$ is their estimated covariance:

$$y_1 = \frac{\left(\sum b_k\right)\left(\sum \cdot_k^2\right) - \left(\sum b_k \cdot_k\right)\left(\sum \cdot_k\right)}{D} \tag{H.13a}$$

$$y_2 = \frac{n\sum b_k \cdot_k - \left(\sum b_k\right)\left(\sum \cdot_k\right)}{D} \tag{H.13b}$$

$$s^2(y_1) = \frac{s^2 \sum \cdot_k^2}{D} \tag{H.13c}$$

$$s^2(y_2) = n\frac{s^2}{D} \tag{H.13d}$$

$$r(y_1, y_2) = -\frac{\sum \cdot_k}{\sqrt{n\sum \cdot_k^2}} \tag{H.13e}$$

$$s^2 = \frac{\sum\left[b_k - b(t_k)\right]^2}{n-2} \tag{H.13f}$$

$$D = n\sum \cdot_k^2 - \left(\sum \cdot_k\right)^2 = n\sum\left(\cdot_k - \bar{\cdot}\right)^2 = n\sum\left(t_k - \bar{t}\right)^2 \tag{H.13g}$$

where all sums are from $k = 1$ to n, $\theta_k = t_k - t_0$, $\bar{\cdot} = \left(\sum \cdot_k\right)/n$, and $\bar{t} = \left(\sum t_k\right)/n$; $[b_k - b(t_k)]$ is the difference between the measured or observed correction b_k at the temperature t_k and the correction $b(t_k)$ predicted by the fitted curve $b(t) = y_1 + y_2(t - t_0)$ at t_k. The variance s^2 is a measure of the overall uncertainty of the fit, where the factor $n - 2$ reflects the fact that because two parameters, y_1 and y_2, are determined by the n observations, the degrees of freedom of s^2 is $\nu = n - 2$ (see G.3.3).

H.3.3 Calculation of results

The data to be fitted are given in the second and third columns of Table H.6. Taking $t_0 = 20$ °C as the reference temperature, application of Equations (H.13a) to (H.13g) yields

$y_1 = -0{,}171\,2$ °C $\quad s(y_1) = 0{,}002\,9$ °C
$y_2 = 0{,}002\,18$ $\quad s(y_2) = 0{,}000\,67$
$r(y_1, y_2) = -0{,}930$ $\quad s = 0{,}003\,5$ °C

The fact that the slope y_2 is more than three times larger than its standard uncertainty provides some indication that a calibration curve and not a fixed average correction is required.

Table H.6 — Data used to obtain a linear calibration curve for a thermometer by the method of least squares

Reading number k	Thermometer reading t_k (°C)	Observed correction $b_k = t_{R,k} - t_k$ (°C)	Predicted correction $b(t_k)$ (°C)	Difference between observed and predicted correction $b_k - b(t_k)$ (°C)
1	21,521	−0,171	−0,167 9	−0,003 1
2	22,012	−0,169	−0,166 8	−0,002 2
3	22,512	−0,166	−0,165 7	−0,000 3
4	23,003	−0,159	−0,164 6	+0,005 6
5	23,507	−0,164	−0,163 5	−0,000 5
6	23,999	−0,165	−0,162 5	−0,002 5
7	24,513	−0,156	−0,161 4	+0,005 4
8	25,002	−0,157	−0,160 3	+0,003 3
9	25,503	−0,159	−0,159 2	+0,000 2
10	26,010	−0,161	−0,158 1	−0,002 9
11	26,511	−0,160	−0,157 0	−0,003 0

The calibration curve may then be written as

$$b(t) = -0{,}171\,2(29)\ °\text{C} + 0{,}002\,18(67)(t - 20\ °\text{C}) \tag{H.14}$$

where the numbers in parentheses are the numerical values of the standard uncertainties referred to the corresponding last digits of the quoted results for the intercept and slope (see 7.2.2). This equation gives the predicted value of the correction $b(t)$ at any temperature t, and in particular the value $b(t_k)$ at $t = t_k$. These values are given in the fourth column of the table while the last column gives the differences between the measured and predicted values, $b_k - b(t_k)$. An analysis of these differences can be used to check the validity of the linear model; formal tests exist (see Reference [8]), but are not considered in this example.

H.3.4 Uncertainty of a predicted value

The expression for the combined standard uncertainty of the predicted value of a correction can be readily obtained by applying the law of propagation of uncertainty, Equation (16) in 5.2.2, to Equation (H.12). Noting that $b(t) = f(y_1, y_2)$ and writing $u(y_1) = s(y_1)$ and $u(y_2) = s(y_2)$, one obtains

$$u_c^2[b(t)] = u^2(y_1) + (t - t_0)^2 u^2(y_2) + 2(t - t_0)\,u(y_1)\,u(y_2)\,r(y_1, y_2) \tag{H.15}$$

The estimated variance $u_c^2[b(t)]$ is a minimum at $t_{min} = t_0 - u(y_1)r(y_1, y_2)/u(y_2)$, which in the present case is t_{min} = 24,008 5 °C.

As an example of the use of Equation (H.15), consider that one requires the thermometer correction and its uncertainty at $t = 30$ °C, which is outside the temperature range in which the thermometer was actually calibrated. Substituting $t = 30$ °C in Equation (H.14) gives

$$b(30\ °\text{C}) = -0{,}149\,4\ °\text{C}$$

while Equation (H.15) becomes

$$u_c^2\left[b(30\ °C)\right] = (0{,}002\,9\ °C)^2 + (10\ °C)^2\,(0{,}000\,67)^2 + 2(10\ °C)(0{,}002\,9\ °C)(0{,}000\,67)(-0{,}930)$$
$$= 17{,}1 \cdot 10^{-6}\ °C^2$$

or

$$u_c\left[b(30\ °C)\right] = 0{,}004\,1\ °C$$

Thus the correction at 30 °C is –0,149 4 °C, with a combined standard uncertainty of $u_c = 0{,}004\,1$ °C, and with u_c having $\nu = n - 2 = 9$ degrees of freedom.

H.3.5 Elimination of the correlation between the slope and intercept

Equation (H.13e) for the correlation coefficient $r(y_1, y_2)$ implies that if t_0 is so chosen that $\sum_{k=1}^{n} c_k = \sum_{k=1}^{n} (t_k - t_0) = 0$, then $r(y_1, y_2) = 0$ and y_1 and y_2 will be uncorrelated, thereby simplifying the computation of the standard uncertainty of a predicted correction. Since $\sum_{k=1}^{n} c_k = 0$ when $t_0 = \bar{t} = \left(\sum_{k=1}^{n} t_k\right)/n$, and $\bar{t} = 24{,}008\,5$ °C in the present case, repeating the least-squares fit with $t_0 = \bar{t} = 24{,}008\,5$ °C would lead to values of y_1 and y_2 that are uncorrelated. (The temperature $\bar{t}$ is also the temperature at which $u^2[b(t)]$ is a minimum — see H.3.4.) However, repeating the fit is unnecessary because it can be shown that

$$b(t) = y_1' + y_2\,(t - \bar{t}) \tag{H.16a}$$

$$u_c^2\left[b(t)\right] = u^2(y_1') + (t - \bar{t})^2\,u^2(y_2) \tag{H.16b}$$

$$r(y_1', y_2) = 0 \tag{H.16c}$$

where

$$y_1' = y_1 + y_2\,(\bar{t} - t_0)$$

$$\bar{t} = t_0 - s(y_1)\,r(y_1, y_2)/s(y_2)$$

$$s^2(y_1') = s^2(y_1)\left[1 - r^2(y_1, y_2)\right]$$

and in writing Equation (H.16b), the substitutions $u(y_1') = s(y_1')$ and $u(y_2) = s(y_2)$ have been made [see Equation (H.15)].

Application of these relations to the results given in H.3.3 yields

$$b(t) = -0{,}162\,5(11) + 0{,}002\,18(67)(t - 24{,}008\,5\ °C) \tag{H.17a}$$

$$u_c^2\left[b(t)\right] = (0{,}001\,1)^2 + (t - 24{,}008\,5\ °C)^2\,(0{,}000\,67)^2 \tag{H.17b}$$

That these expressions give the same results as Equations (H.14) and (H.15) can be checked by repeating the calculation of b(30 °C) and $u_c[b(30\ °C)]$. The substitution of $t = 30$ °C into Equations (H.17a) and (H.17b) yields

$$b(30\ °C) = -0{,}149\,4\ °C$$

$$u_c\left[b(30\ °C)\right] = 0{,}004\,1\ °C$$

which are identical to the results obtained in H.3.4. The estimated covariance between two predicted corrections $b(t_1)$ and $b(t_2)$ may be obtained from Equation (H.9) in H.2.3.

H.3.6 Other considerations

The least-squares method can be used to fit higher-order curves to data points, and is also applicable to cases where the individual data points have uncertainties. Standard texts on the subject should be consulted for details [8]. However, the following examples illustrate two cases where the measured corrections b_k are not assumed to be exactly known.

1) Let each t_k have negligible uncertainty, let each of the n values $t_{\mathrm{R},\,k}$ be obtained from a series of m repeated readings, and let the pooled estimate of variance for such readings based on a large amount of data obtained over several months be s_{p}^2. Then the estimated variance of each $t_{\mathrm{R},\,k}$ is $s_{\mathrm{p}}^2/m = u_0^2$ and each observed correction $b_k = t_{\mathrm{R},\,k} - t_k$ has the *same* standard uncertainty u_0. Under these circumstances (and under the assumption that there is no reason to believe that the linear model is incorrect), u_0^2 replaces s^2 in Equations (H.13c) and (H.13d).

   NOTE A pooled estimate of variance s_{p}^2 based on N series of independent observations of the same random variable is obtained from

$$s_{\mathrm{p}}^2 = \frac{\sum_{i=1}^{N} \nu_i s_i^2}{\sum_{i=1}^{N} \nu_i}$$

   where s_i^2 is the experimental variance of the ith series of n_i independent repeated observations [Equation (4) in 4.2.2] and has degrees of freedom $\nu_i = n_i - 1$. The degrees of freedom of s_{p}^2 is $\nu = \sum_{i=1}^{N} \nu_i$. The experimental variance s_{p}^2/m (and the experimental standard deviation $s_{\mathrm{p}}/\sqrt{m}$) of the arithmetic mean of m independent observations characterized by the pooled estimate of variance s_{p}^2 also has ν degrees of freedom.

2) Suppose that each t_k has negligible uncertainty, that a correction ε_k is applied to each of the n values $t_{\mathrm{R},\,k}$, and that each correction has the same standard uncertainty u_{a}. Then the standard uncertainty of each $b_k = t_{\mathrm{R},\,k} - t_k$ is also u_{a}, and $s^2(y_1)$ is replaced by $s^2(y_1) + u_{\mathrm{a}}^2$ and $s^2(y_2)$ is replaced by $s^2(y_2) + u_{\mathrm{a}}^2$.

H.4 Measurement of activity

This example is similar to example H.2, the simultaneous measurement of resistance and reactance, in that the data can be analysed in two different ways but each yields essentially the same numerical result. The first approach illustrates once again the need to take the observed correlations between input quantities into account.

H.4.1 The measurement problem

The unknown radon (^{222}Rn) activity concentration in a water sample is determined by liquid-scintillation counting against a radon-in-water standard sample having a known activity concentration. The unknown activity concentration is obtained by measuring three counting sources consisting of approximately 5 g of water and 12 g of organic emulsion scintillator in vials of volume 22 ml:

Source (a) a *standard* consisting of a mass m_{S} of the standard solution with a known activity concentration;

Source (b) a matched *blank* water sample containing no radioactive material, used to obtain the background counting rate;

Source (c) the *sample* consisting of an aliquot of mass m_x with unknown activity concentration.

Six cycles of measurement of the three counting sources are made in the order standard — blank — sample; and each dead-time-corrected counting interval T_0 for each source during all six cycles is 60 minutes. Although the background counting rate cannot be assumed to be constant over the entire counting interval (65 hours), it is assumed that the number of counts obtained for each blank may be used as representative of the background counting rate during the measurements of the standard and sample in the same cycle. The data are given in Table H.7, where

t_S, t_B, t_x are the times from the reference time $t = 0$ to the midpoint of the dead-time-corrected counting intervals $T_0 = 60$ min for the standard, blank, and sample vials, respectively; although t_B is given for completeness, it is not needed in the analysis;

C_S, C_B, C_x are the number of counts recorded in the dead-time-corrected counting intervals $T_0 = 60$ min for the standard, blank, and sample vials, respectively.

The observed counts may be expressed as

$$C_S = C_B + \varepsilon A_S T_0 m_S e^{-\cdot t_S} \quad \text{(H.18a)}$$

$$C_x = C_B + \varepsilon A_x T_0 m_x e^{-\cdot t_x} \quad \text{(H.18b)}$$

where

ε is the liquid scintillation detection efficiency for ^{222}Rn for a given source composition, assumed to be independent of the activity level;

A_S is the activity concentration of the standard at the reference time $t = 0$;

A_x is the *measurand* and is defined as the unknown activity concentration of the sample at the reference time $t = 0$;

m_S is the mass of the standard solution;

m_x is the mass of the sample aliquot;

. is the decay constant for ^{222}Rn: . $= (\ln 2)/T_{1/2} = 1{,}258\ 94 \cdot 10^{-4}$ min^{-1} ($T_{1/2} = 5\ 505{,}8$ min).

Table H.7 — Counting data for determining the activity concentration of an unknown sample

Cycle	Standard		Blank		Sample	
k	t_S (min)	C_S (counts)	t_B (min)	C_B (counts)	t_x (min)	C_x (counts)
1	243,74	15 380	305,56	4 054	367,37	41 432
2	984,53	14 978	1 046,10	3 922	1 107,66	38 706
3	1 723,87	14 394	1 785,43	4 200	1 846,99	35 860
4	2 463,17	13 254	2 524,73	3 830	2 586,28	32 238
5	3 217,56	12 516	3 279,12	3 956	3 340,68	29 640
6	3 956,83	11 058	4 018,38	3 980	4 079,94	26 356

Equations (H.18a) and (H.18b) indicate that neither the six individual values of C_S nor of C_x given in Table H.7 can be averaged directly because of the exponential decay of the activity of the standard and sample, and slight variations in background counts from one cycle to another. Instead, one must deal with the decay-corrected and background-corrected counts (or counting rates defined as the number of counts divided by $T_0 = 60$ min). This suggests combining Equations (H.18a) and (H.18b) to obtain the following expression for the unknown concentration in terms of the known quantities:

$$\begin{aligned} A_x &= f\left(A_S, m_S, m_x, C_S, C_x, C_B, t_S, t_x, .\right) \\ &= A_S \frac{m_S}{m_x} \frac{\left(C_x - C_B\right)\mathrm{e}^{\cdot\, t_x}}{\left(C_S - C_B\right)\mathrm{e}^{\cdot\, t_S}} \\ &= A_S \frac{m_S}{m_x} \frac{C_x - C_B}{C_S - C_B} \mathrm{e}^{\cdot\, (t_x - t_S)} \end{aligned} \tag{H.19}$$

where $(C_x - C_B)\mathrm{e}^{\cdot\, t_x}$ and $(C_S - C_B)\mathrm{e}^{\cdot\, t_S}$ are, respectively, the background-corrected counts of the sample and the standard at the reference time $t = 0$ and for the time interval $T_0 = 60$ min. Alternatively, one may simply write

$$A_x = f\left(A_S, m_S, m_x, R_S, R_x\right) = A_S \frac{m_S}{m_x} \frac{R_x}{R_S} \tag{H.20}$$

where the background-corrected and decay-corrected *counting rates* R_x and R_S are given by

$$R_x = \left[\left(C_x - C_B\right)/T_0\right]\mathrm{e}^{\cdot\, t_x} \tag{H.21a}$$

$$R_S = \left[\left(C_S - C_B\right)/T_0\right]\mathrm{e}^{\cdot\, t_S} \tag{H.21b}$$

H.4.2 Analysis of data

Table H.8 summarizes the values of the background-corrected and decay-corrected counting rates R_S and R_x calculated from Equations (H.21a) and (H.21b) using the data of Table H.7 and . = 1,258 94 · 10^{-4} min^{-1} as given earlier. It should be noted that the ratio $R = R_x/R_S$ is most simply calculated from the expression

$$\left[\left(C_x - C_B\right)/\left(C_S - C_B\right)\right]\mathrm{e}^{\cdot\, (t_x - t_S)}$$

The arithmetic means $\bar{R}_S$, $\bar{R}_x$, and $\bar{R}$, and their experimental standard deviations $s(\bar{R}_S)$, $s(\bar{R}_x)$, and $s(\bar{R})$, are calculated in the usual way [Equations (3) and (5) in 4.2]. The correlation coefficient $r(\bar{R}_x, \bar{R}_S)$ is calculated from Equation (17) in 5.2.3 and Equation (14) in 5.2.2.

Because of the comparatively small variability of the values of R_x and of R_S, the ratio of means $\bar{R}_x / \bar{R}_S$ and the standard uncertainty $u(\bar{R}_x / \bar{R}_S)$ of this ratio are, respectively, very nearly the same as the mean ratio $\bar{R}$ and its experimental standard deviation $s(\bar{R})$ as given in the last column of Table H.8 [see H.2.4 and Equation (H.10) therein]. However, in calculating the standard uncertainty $u(\bar{R}_x / \bar{R}_S)$, the correlation between R_x and R_S as represented by the correlation coefficient $r(\bar{R}_x, \bar{R}_S)$ must be taken into account using Equation (16) in 5.2.2. [That equation yields for the relative estimated variance of $\bar{R}_x / \bar{R}_S$ the last three terms of Equation (H.22b).]

It should be recognized that the respective experimental standard deviations of R_x and of R_S, $\sqrt{6}\, s(\bar{R}_x)$ and $\sqrt{6}\, s(\bar{R}_S)$, indicate a variability in these quantities that is two to three times larger than the variability implied by the Poisson statistics of the counting process; the latter is included in the observed variability of the counts and need not be accounted for separately.

Table H.8 — Calculation of decay-corrected and background-corrected counting rates

Cycle k	R_x (min^{-1})	R_S (min^{-1})	$t_x - t_S$ (min)	$R = R_x/R_S$
1	652,46	194,65	123,63	3,352 0
2	666,48	208,58	123,13	3,195 3
3	665,80	211,08	123,12	3,154 3
4	655,68	214,17	123,11	3,061 5
5	651,87	213,92	123,12	3,047 3
6	623,31	194,13	123,11	3,210 7
	$\bar{R}_x = 652{,}60$ $s(\bar{R}_x) = 6{,}42$ $s(\bar{R}_x)/\bar{R}_x = 0{,}98 \cdot 10^{-2}$	$\bar{R}_S = 206{,}09$ $s(\bar{R}_S) = 3{,}79$ $s(\bar{R}_S)/\bar{R}_S = 1{,}84 \cdot 10^{-2}$		$\bar{R} = 3{,}170$ $s(\bar{R}) = 0{,}046$ $s(\bar{R})/\bar{R} = 1{,}44 \cdot 10^{-2}$
	$\bar{R}_x/\bar{R}_S = 3{,}167$ $u(\bar{R}_x/\bar{R}_S) = 0{,}045$ $u(\bar{R}_x/\bar{R}_S)/(\bar{R}_x/\bar{R}_S) = 1{,}42 \cdot 10^{-2}$			
Correlation coefficient				
$r(\bar{R}_x, \bar{R}_S) = 0{,}646$				

H.4.3 Calculation of final results

To obtain the unknown activity concentration A_x and its combined standard uncertainty $u_c(A_x)$ from Equation (H.20) requires A_S, m_x, and m_S and their standard uncertainties. These are given as

$A_S = 0{,}136\ 8$ Bq/g
$u(A_S) = 0{,}001\ 8$ Bq/g; $\quad u(A_S)/A_S = 1{,}32 \cdot 10^{-2}$

$m_S = 5{,}019\ 2$ g
$u(m_S) = 0{,}005\ 0$ g; $\quad u(m_S)/m_S = 0{,}10 \cdot 10^{-2}$

$m_x = 5{,}057\ 1$ g
$u(m_x) = 0{,}001\ 0$ g; $\quad u(m_x)/m_x = 0{,}02 \cdot 10^{-2}$

Other possible sources of uncertainty are evaluated to be negligible:

- standard uncertainties of the decay times, $u(t_{S,k})$ and $u(t_{x,k})$;
- standard uncertainty of the decay constant of ^{222}Rn, $u(\,.\,) = 1 \cdot 10^{-7}$ min^{-1}. (The significant quantity is the decay factor exp[. $(t_x - t_S)$], which varies from 1,015 63 for cycles k = 4 and 6 to 1,015 70 for cycle k = 1. The standard uncertainty of these values is $u = 1{,}2 \cdot 10^{-5}$);
- uncertainty associated with the possible dependence of the detection efficiency of the scintillation counter on the source used (standard, blank, and sample);
- uncertainty of the correction for counter dead-time and of the correction for the dependence of counting efficiency on activity level.