Essentials of expressing measurement uncertainty

This is a brief summary of the method of evaluating and expressing uncertainty in measurement adopted widely by U.S. industry, companies in other countries, NIST, its sister national metrology institutes throughout the world, and many organizations worldwide. These "essentials" are adapted from NIST Technical Note 1297 (TN 1297), prepared by B.N. Taylor and C.E. Kuyatt and entitled *Guidelines for Evaluating and Expressing the Uncertainty of NIST Measurement Results*, which in turn is based on the comprehensive International Organization for Standardization (ISO) *Guide to the Expression of Uncertainty in Measurement*. Users requiring more detailed information may see TN 1297, or if a comprehensive discussion is desired, they may purchase the ISO *Guide*.

Background information on the development of the ISO *Guide*, its worldwide adoption, NIST TN 1297, and the NIST policy on expressing measurement uncertainty is given in the section **International and U.S. perspectives on measurement uncertainty**.

To assist you in reading these guidelines, you may wish to consult a short **glossary**. Additionally, a companion publication to the ISO *Guide,* entitled the *International Vocabulary of Basic and General Terms in Metrology*, or VIM, gives definitions of many other important terms relevant to the field of measurement.

Basic definitions

Measurement equation

The case of interest is where the quantity Y being measured, called the **measurand**, is not measured directly, but is determined from N other quantities $X_1, X_2, \ldots, X_N$ through a functional relation f, often called the **measurement equation**:

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

Included among the quantities X_i are corrections (or correction factors), as well as quantities that take into account other sources of variability, such as different observers, instruments, samples, laboratories, and times at which observations are made (e.g., different days). Thus, the function f of equation (1) should express not simply a physical law but a measurement process, and in particular, it should contain all quantities that can contribute a significant uncertainty to the measurement result.

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

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

For example, as pointed out in the ISO *Guide*, 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 b, the power P (the measurand) dissipated by the resistor at the temperature t depends on V, R_0, b, and t according to

$$P = f(V, R_0, b, t) = V^2/R_0[1 + b(t - t_0)]. \tag{3}$$

Classification of uncertainty components

The uncertainty of the measurement result y arises from the uncertainties $u\,(x_i)$ (or u_i for brevity) of the input estimates x_i that enter equation (2). Thus, in the example of equation (3), the uncertainty of the estimated value of the power P arises from the uncertainties of the estimated values of the potential difference V, resistance R_0, temperature coefficient of resistance b, and temperature t. In general, components of uncertainty may be categorized according to the **method** used to evaluate them.

> **Type A evaluation**
> method of evaluation of uncertainty by the **statistical analysis** of series of observations,
>
> **Type B evaluation**
> method of evaluation of uncertainty by means **other than the statistical analysis** of series of observations.

Representation of uncertainty components

Standard Uncertainty
Each component of uncertainty, however evaluated, is represented by an estimated standard deviation, termed **standard uncertainty** with suggested symbol u_i, and equal to the positive square root of the estimated variance

Standard uncertainty: Type A
An uncertainty component obtained by a Type A evaluation is represented by a statistically estimated standard deviation s_i, equal to the positive square root of the statistically estimated variance s_i^2, and the associated number of degrees of freedom v_i. For such a component the standard uncertainty is $u_i = s_i$.

Standard uncertainty: Type B
In a similar manner, an uncertainty component obtained by a Type B evaluation is represented by a quantity u_j , which may be considered an approximation to the corresponding standard deviation; it is equal to the positive square root of u_j^2, which may be considered an approximation to the corresponding variance and which is obtained from an assumed probability distribution based on all the available information. Since the quantity u_j^2 is treated like a variance and u_j like a standard deviation, for such a component the standard uncertainty is simply u_j.

Evaluating uncertainty components: Type A

A Type A evaluation of standard uncertainty may be based on any valid statistical method for treating data. Examples are calculating the standard deviation of the mean of a series of independent observations; using the method of least squares to fit a curve to data in order to estimate the parameters of the curve and their standard deviations; and carrying out an analysis of variance (ANOVA) in order to identify and quantify random effects in certain kinds of measurements.

Mean and standard deviation

As an example of a Type A evaluation, consider an input quantity X_i whose value is estimated from n independent observations $X_{i,k}$ of X_i obtained under the same conditions of measurement. In this case the input estimate x_i is usually the **sample mean**

$$x_i = \bar{X}_i = \frac{1}{n}\sum_{k=1}^{n} X_{i,k} \qquad (4)$$

and the standard uncertainty $u(x_i)$ to be associated with x_i is the estimated **standard deviation of the mean**

$$u(x_i) = s(\bar{X}_i) = \left(\frac{1}{n(n-1)}\sum_{k=1}^{n}(X_{i,k}-\bar{X}_i)^2\right)^{1/2} \qquad (5)$$

Evaluating uncertainty components: Type B

A Type B evaluation of standard uncertainty is usually based on scientific judgment using all of the relevant information available, which may include:

- previous measurement data,
- experience with, or general knowledge of, the behavior and property of relevant materials and instruments,
- manufacturer's specifications,
- data provided in calibration and other reports, and
- uncertainties assigned to reference data taken from handbooks.

Below are some examples of Type B evaluations in different situations, depending on the available information and the assumptions of the experimenter. Broadly speaking, the uncertainty is either obtained from an outside source, or obtained from an assumed distribution.

Uncertainty obtained from an outside source

Multiple of a standard deviation

Procedure: Convert an uncertainty quoted in a handbook, manufacturer's specification, calibration certificate, etc., that is a stated multiple of an estimated standard deviation to a standard uncertainty by dividing the quoted uncertainty by the multiplier.

Confidence interval

Procedure: Convert an uncertainty quoted in a handbook, manufacturer's specification, calibration certificate, etc., that defines a "confidence interval" having a stated level of confidence, such as 95 % or 99 %, to a standard uncertainty by treating the quoted uncertainty as if a normal probability distribution had been used to calculate it (unless otherwise indicated) and dividing it by the appropriate factor for such a distribution. These factors are 1.960 and 2.576 for the two levels of confidence given.

Uncertainty obtained from an assumed distribution

Normal distribution: "1 out of 2"

Procedure: Model the input quantity in question by a normal probability distribution and estimate lower and upper limits a_- and a_+ such that the best estimated value of the input quantity is $(a_+ + a_-)/2$ (i.e., the center of the limits) and there is 1 chance out of 2 (i.e., a 50 % probability) that the value of the quantity lies in the interval a_- to a_+. Then u_j is approximately 1.48 a, where $a = (a_+ - a_-)/2$ is the half-width of the interval.

Normal distribution: "2 out of 3"

Procedure: Model the input quantity in question by a normal probability distribution and estimate lower and upper limits a_- and a_+ such that the best estimated value of the input quantity is $(a_+ + a_-)/2$ (i.e., the center of the limits) and there are 2 chances out of 3 (i.e., a 67 % probability) that the value of the quantity lies in the interval a_- to a_+. Then u_j is approximately a, where $a = (a_+ - a_-)/2$ is the half-width of the interval.

Normal distribution: "99.73 %"

Procedure: If the quantity in question is modeled by a normal probability distribution, there are no finite limits that will contain 100 % of its possible values. However, plus and minus 3 standard deviations about the mean of a normal distribution corresponds to 99.73 % limits. Thus, if the limits a_- and a_+ of a normally distributed quantity with mean $(a_+ + a_-)/2$ are considered to contain "almost all" of the possible values of the quantity, that is, approximately 99.73 % of them, then u_j is approximately $a/3$, where $a = (a_+ - a_-)/2$ is the half-width of the interval.

Uniform (rectangular) distribution

Procedure: Estimate lower and upper limits a_- and a_+ for the value of the input quantity in question such that the probability that the value lies in the interval a_- and a_+ is, for all practical purposes, 100 %. Provided that there is no contradictory information, treat the quantity as if it is equally probable for its value to lie anywhere within the interval a_- to a_+; that is, model it by a uniform (i.e., rectangular) probability distribution. The best estimate of the value of the quantity is then $(a_+ + a_-)/2$ with $u_j = a$ divided by the square root of 3, where $a = (a_+ - a_-)/2$ is the half-width of the interval.

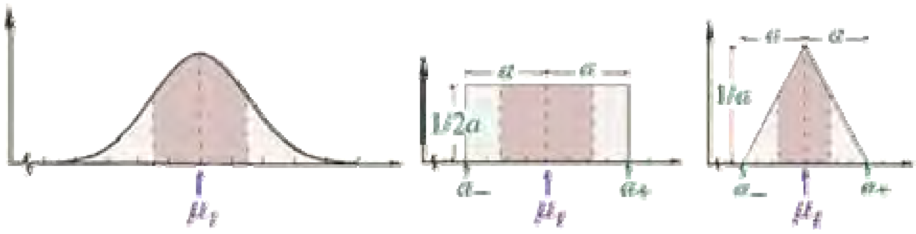
Triangular distribution

The rectangular distribution is a reasonable default model in the absence of any other information. But if it is known that values of the quantity in question near the center of the limits are more likely than values close to the limits, a normal distribution or, for simplicity, a triangular distribution, may be a better model.

Procedure: Estimate lower and upper limits a_- and a_+ for the value of the input quantity in question such that the probability that the value lies in the interval a_- to a_+ is, for all practical purposes, 100 %. Provided that there is no contradictory information, model the quantity by a triangular probability distribution. The best estimate of the value of the quantity is then $(a_+ + a_-)/2$ with $u_j = a$ divided by the square root of 6, where $a = (a_+ - a_-)/2$ is the half-width of the interval.

Schematic illustration of probability distributions

The following figure schematically illustrates the three distributions described above: normal, rectangular, and triangular. In the figures, μ_t is the expectation or mean of the distribution, and the shaded areas represent ± one standard uncertainty u about the mean. For a normal distribution, $\pm u$ encompases about 68 % of the distribution; for a uniform distribution, $\pm u$ encompasses about 58 % of the distribution; and for a triangular distribution, $\pm u$ encompasses about 65 % of the distribution.



Combining uncertainty components

Calculation of combined standard uncertainty

The **combined standard uncertainty** of the measurement result y, designated by $u_c(y)$ and taken to represent the estimated standard deviation of the result, is the positive square root of the estimated variance $u_c^2(y)$ obtained from

$$u_c^2(y) = \sum_{i=1}^{N} \left(\frac{\partial f}{\partial x_i}\right)^2 u^2(x_i) + 2\sum_{i=1}^{N-1} \sum_{j=i+1}^{N} \frac{\partial f}{\partial x_i}\frac{\partial f}{\partial x_j} u(x_i, x_j) \qquad (6)$$

Equation (6) is based on a first-order Taylor series approximation of the measurement equation $Y = f(X_1, X_2, \ldots, X_N)$ given in equation (1) and is conveniently referred to as the *law of propagation of uncertainty*. The partial derivatives of f with respect to the X_i (often referred to as *sensitivity coefficients*) are equal to the partial derivatives of f with respect to the X_i evaluated at $X_i = x_i$; $u(x_i)$ is the standard uncertainty associated with the input estimate x_i; and $u(x_i, x_j)$ is the estimated covariance associated with x_i and x_j.

Simplified forms

Equation (6) often reduces to a simple form in cases of practical interest. For example, if the input estimates x_i of the input quantities X_i can be assumed to be uncorrelated, then the second term vanishes. Further, if the input estimates are uncorrelated and the measurement equation is one of the following two forms, then equation (6) becomes simpler still.

Measurement equation:

A sum of quantities X_i multiplied by constants a_i.

$$Y = a_1X_1 + a_2X_2 + \ldots a_NX_N$$

Measurement result:

$$y = a_1x_1 + a_2x_2 + \ldots a_Nx_N$$

Combined standard uncertainty:

$$u_c^2(y) = a_1^2u^2(x_1) + a_2^2u^2(x_2) + \ldots a_N^2u^2(x_N)$$

Measurement equation:

A product of quantities X_i, raised to powers $a, b, \ldots p$, multiplied by a constant A.

$$Y = AX_1^{\,a}X_2^{\,b} \ldots X_N^{\,p}$$

Measurement result:

$$y = Ax_1^{\,a}x_2^{\,b} \ldots x_N^{\,p}$$

Combined standard uncertainty:

$$u_{c,r}^2(y) = a^2u_r^2(x_1) + b^2u_r^2(x_2) + \ldots p^2u_r^2(x_N)$$

Here $u_r(x_i)$ is the **relative standard uncertainty** of x_i and is defined by $u_r(x_i) = u(x_i)/|x_i|$, where $|x_i|$ is the absolute value of x_i and x_i is not equal to zero; and $u_{c,r}(y)$ is the **relative combined standard uncertainty** of y and is defined by $u_{c,r}(y) = u_c(y)/|y|$, where $|y|$ is the absolute value of y and y is not equal to zero.

Meaning of uncertainty

If the probability distribution characterized by the measurement result y and its combined standard uncertainty $u_c(y)$ is approximately normal (Gaussian), and $u_c(y)$ is a reliable estimate of the standard deviation of y, then the interval y $u_c(y)$ to $y + u_c(y)$ is expected to encompass approximately 68 % of the distribution of values that could reasonably be attributed to the value of the quantity Y of which y is an estimate. This implies that it is believed with an approximate level of confidence of 68 % that Y is greater than or equal to y $u_c(y)$, and is less than or equal to $y + u_c(y)$, which is commonly written as $Y = y \pm u_c(y)$.

Expanded uncertainty and coverage factor

Expanded uncertainty

Although the combined standard uncertainty u_c is used to express the uncertainty of many measurement results, for some commercial, industrial, and regulatory applications (e.g., when health and safety are concerned), what is often required is a measure of uncertainty that defines an interval about the measurement result y within which the value of the measurand Y can be confidently asserted to lie. The measure of uncertainty intended to meet this requirement is termed **expanded uncertainty**, suggested symbol U, and is obtained by multiplying $u_c(y)$ by a **coverage factor**, suggested symbol k. Thus $U = ku_c(y)$ and it is confidently believed that Y is greater than or equal to $y - U$, and is less than or equal to $y + U$, which is commonly written as $Y = y \pm U$.

Coverage factor

In general, the value of the coverage factor k is chosen on the basis of the desired level of confidence to be associated with the interval defined by $U = ku_c$. Typically, k is in the range 2 to

3. When the normal distribution applies and u_c is a reliable estimate of the standard deviation of y, $U = 2\ u_c$ (i.e., $k = 2$) defines an interval having a level of confidence of approximately 95 %, and $U = 3\ u_c$ (i.e., $k = 3$) defines an interval having a level of confidence greater than 99 %.

Relative expanded uncertainty

In analogy with relative standard uncertainty u_r and relative combined standard uncertainty $u_{c,r}$ defined above in connection with simplified forms of equation (6), the **relative expanded uncertainty** of a measurement result y is $U_r = U/|y|$, y not equal to zero.

Examples of uncertainty statements

The following are examples of uncertainty statements as would be used in publication or correspondence. In each case, the quantity whose value is being reported is assumed to be a nominal 100 g standard of mass m_s.

Example 1

m_s = 100.021 47 g with a combined standard uncertainty (i.e., estimated standard deviation) of u_c = 0.35 mg. Since it can be assumed that the possible estimated values of the standard are approximately normally distributed with approximate standard deviation u_c, the unknown value of the standard is believed to lie in the interval $m_s \pm u_c$ with a level of confidence of approximately 68 %.

Example 2

m_s = (100.021 47 ± 0.000 70) g, where the number following the symbol ± is the numerical value of an expanded uncertainty $U = k\ u_c$, with U determined from a combined standard uncertainty (i.e., estimated standard deviation) u_c = 0.35 mg and a coverage factor k = 2. Since it can be assumed that the possible estimated values of the standard are approximately normally distributed with approximate standard deviation u_c, the unknown value of the standard is believed to lie in the interval defined by U with a level of confidence of approximately 95 %.

Background

A measurement result is complete only when accompanied by a quantitative statement of its uncertainty. The uncertainty is required in order to decide if the result is adequate for its intended purpose and to ascertain if it is consistent with other similar results.

International and U.S. perspectives on measurement uncertainty

Over the years, many different approaches to evaluating and expressing the uncertainty of measurement results have been used. Because of this lack of international agreement on the expression of uncertainty in measurement, in 1977 the International Committee for Weights and Measures (CIPM, *Comité International des Poids et Measures*), the world's highest authority in the field of measurement science (*i.e.*, metrology), asked the International Bureau of Weights and Measures (BIPM, *Bureau International des Poids et Mesures*), to address the problem in collaboration with the various national metrology institutes and to propose a specific recommendation for its solution. This led to the development of Recommendation INC-1 (1980) by the Working Group on the Statement of Uncertainties convened by the BIPM, a recommendation that the CIPM approved in 1981 and reaffirmed in 1986 via its own Recommendations 1 (CI-1981) and 1 (CI-1986):

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.

Type A. Those which are evaluated by statistical methods

Type 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 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 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 an overall uncertainty, the multiplying factor must always be stated.

The above recommendation, INC-1 (1980), is a brief outline rather than a detailed prescription. Consequently, the CIPM asked the International Organization for Standardization (ISO) to develop a detailed guide based on the recommendation because ISO could more easily reflect the requirements stemming from the broad interests of industry and commerce. The ISO Technical Advisory Group on Metrology (TAG 4) was given this responsibility. It in turn established Working group 3 and assigned it 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.

Glossary

The following definitions are given in the ISO *Guide to the Expression of Uncertainty in Measurement*. Many additional terms relevant to the field of measurement are given in a companion publication to the ISO *Guide,* entitled the *International Vocabulary of Basic and General Terms in Metrology*, or VIM. Both the ISO *Guide* and VIM

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

- 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.
- Uncertainty of measurement comprises, in general, many components. Some of these components may be evaluated from the statistical distribution of the results of a 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.
- 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.

Standard uncertaintyuncertainty of the result of a measurement expressed as a standard deviation

Type A evaluation (of uncertainty) method of evaluation of uncertainty by the statistical analysis of series of observations

Type B evaluation (of uncertainty) method of evaluation of uncertainty by means other than the statistical analysis of series of observations

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 weighed according to how the measurement result varies with changes in these quantities

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.

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

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