

The Drongo's Guide to Determining Measurement Uncertainty

1. Estimating uncertainty

Uncertainty estimation is not complex or difficult in principle. While some measurement processes are complex and hence the uncertainty calculations take on a degree of complexity, the overall principles remain the same. The International Standards Organization's "Guide to the Expression of Uncertainty in Measurement" (GUM) gives a very comprehensive treatment on the process. Briefly the process is as follows:

- Develop a model of the measurement
- List all the uncertainty components
- Calculate the standard uncertainties.
- Calculate the sensitivity coefficients.
- Calculate the associated degrees of freedom if required.
- Convert all uncertainties into uncertainties in the measurand.
- Combine all the uncertainties.

The GUM states that all standard uncertainties are equivalent to standard deviations.

2. Modeling the measurement system

Before any calculations are attempted it is necessary to consider the measurement system and its environment. A simple sketch or an equation may be all that is required. The model should provide a simple means of describing the relationship between the input parameters and influence quantities to the measurand. This is the step that many metrologists find most difficult. Worked examples should assist in understanding this task. Without a model some significant uncertainties may be overlooked or it may be difficult to determine the sensitivity coefficients.

3. Determining standard uncertainties

There are two types of uncertainties, Type A and Type B. Uncertainties calculated by statistical analysis are all Type A. Uncertainties obtained by estimation, worst case calculations or taken from reports and references or found by any other means are Type B uncertainties. The classification is independent of the distribution type.

For type A assessments it is necessary to choose the appropriate statistical method. For the situation where the average of several measurements is taken, the procedure is to calculate the mean (or average) of the set of values and then calculate a standard deviation known as the *Experimental Standard Deviation of the Mean*. When data is treated to find a "best fit," such as when a curve is fitted, the standard techniques will also yield a standard uncertainty. Type A assessments are perhaps the most straightforward to make.

For type B assessments the knowledge and understanding of the metrologist is crucial. Each component must be examined to determine its limits, that is the range of dispersion, and the nature of the dispersion. This means that it must be decided if the uncertainty component is expected to cluster near a central value or is evenly dispersed across its range, or if it is more

The Drongo's Guide to Determining Measurement Uncertainty

likely to be at the extremes. This analysis is called determining the distribution curve of the uncertainty. Once the range and nature of the dispersion have been decided, specific formulas are applied to find the standard uncertainty component.

The GUM assumes that for each uncertainty component there is a parameter, similar to a variance, which describes the magnitude of the component. While the variances are the key element from a mathematician's view, metrologists prefer to work with standard deviations because they have the same units as the associated parameter. These standard deviations are called standard uncertainties. A measured temperature may have a value of 427 °C with a standard uncertainty of 1.7 °C. This is easier to visualize than a variance of (3 °C)². There are rules that may be used to calculate the standard uncertainties.

4. Type A assessment

When a measurement is repeated several times, the mean value and the standard deviation can be calculated. The standard deviation describes the dispersion applicable to the whole population of possible measured values. The relevant formulas are discussed in the following sections.

4.1 The mean

The mean of a set of repeated measurements is the best estimate of the true mean of the population. It is calculated by dividing the sum of all measurements by n , the number of measurements.

4.2 Standard deviation

The standard deviation (symbol s) is a measure of the dispersion of the population from which the n values were taken. It is readily calculated using a calculator with statistical functions. Care needs to be taken that the button labeled “ s ” or s_{n-1} ” is used to obtain the unweighted estimate of the population standard deviation. In EXCEL[®] the formula STDEV() is used.

4.3 The ESDM

The standard deviation of the mean of a set of measurements, the ESDM, is simply the estimated standard deviation of the population divided by the square root of the number of measurements n .

Having taken a mean of several measurements we expect that this will be closer to the population mean or true mean than any single measurement. Also, although sets of means have a dispersion, we expect a smaller dispersion than for the original or parent population. It follows that we also expect that the mean should have its own standard deviation and that this will be less than that of the parent population. This is correct, and the associated deviation is called the *Experimental Standard Deviation of the Mean*, or *ESDM*.

When the measurand is obtained from a mean of repeated measurements, the *ESDM* is also the standard uncertainty for the dispersion of the measurand associated with random effects. Thus, this type A method of calculation of

The Drongo's Guide to Determining Measurement Uncertainty

uncertainty components involves calculating the *ESDM*. The *ESDM* is given by the formula:

$$ESDM = \frac{s}{\sqrt{n}} \quad (1)$$

where the terms have the same meaning as before.

This shows that increasing the number of repeated measurements reduces the dispersion to be associated with the mean. Unfortunately, once n is large enough to give a good estimate of s , to halve the *ESDM* requires that there be four times as many measurements. It is important to note that, regardless of the shape of the distribution of the population from which the individual measurements have come, the mean belongs to a normal distribution.

5. Type B assessment

When it comes to Type B uncertainty component assessment, the ISO GUM advises that there is "no substitute for critical thinking, intellectual honesty, and professional skill". While this is intended to apply to the whole process it is of most significance when making type B estimates. The simplest example is that of using the uncertainty of a standard which has a calibration certificate. To obtain the standard uncertainty the expanded uncertainty on the certificate is divided by the coverage factor given on the certificate.

In the absence of a value for the coverage factor, a factor of 2 may be used if the expanded uncertainty has a 95% confidence level.

Thus:

$$u(x_i) = \frac{U_{95\%}}{2} \quad (2)$$

where: $u(x_i)$ is the standard uncertainty of the calibration values given on the certificate and

$U_{95\%}$ is the expanded uncertainty at a 95% confidence limit given on the certificate.

Equation (2) gives an approximate value for $u(x_i)$. When k , the coverage factor is known it should be used instead of the factor 2.

In other cases the uncertainty can usually be given limits, $\pm a$ and some distribution determined. The formula for converting a a to a standard uncertainty depends on the distribution type.

5.1 Normal distribution

For many uncertainties the appropriate distribution is a normal distribution. Given the semi-range, a , it is necessary to know, at least approximately what percentage of the total distribution is encompassed by the range. A table for the Normal distribution can be used to determine the number of standard deviations this represents. In the previous section the range is assumed to cover 95% of the total, so a divisor of 2 is used. If it had represented 68% then a divisor of 1 would apply and if it represented 99.7% the divisor would have been 3.

5.2 Rectangular distribution

If limits can be determined, but the value of the measurand is just as likely to be anywhere in the range, then the distribution of the uncertainty is a rectangular distribution. It is sometimes called the distribution of minimum knowledge, and is most often applied when worst case limits are calculated.

For example, take the case where the laboratory temperature was known to have a range of 19.5 °C to 20.5 °C. This means we have $a = 0.5$ °C. If we had no further information, then we can only say that it is equally likely that the temperature can have any value in that range. This is a uniform or rectangular distribution, and the formula for the standard deviation of such a distribution gives us the standard uncertainty due to the variation in temperature.

The formula is:

$$u(x_i) = \frac{a}{\sqrt{3}} \quad (3)$$

where: $u(x_i)$ is the standard uncertainty and

a is the semi-range of the limits of the uncertainty component.

There are other formulas for all other distributions that may be appropriate. These distributions can be used if we have more knowledge of the nature of the dispersion.

6. Sensitivity coefficients

This is the other aspect of uncertainty estimation that may cause some difficulty. The sensitivity coefficient is the factor that converts an uncertainty component to the same units as the measurand.

This is a necessary precondition to combining the components as combining, say, uncertainties in metres with uncertainties in kilograms is not dimensionally correct. Common sense would also tell us that only like items should be combined.

The sensitivity coefficient is a measure of the sensitivity of the measurand to a particular input or influence parameter. For example, a steel tape has roughly twelve times the linear temperature expansion coefficient of an invar tape (11.5 ppm/°C compared to 1 ppm/°C). Thus, for measurements involving a steel tape, the measured length is more sensitive to variations in the tape temperature than for the invar one. So the sensitivity of the uncertainty of the measured length to temperature is greater for steel tapes than for invar tapes.

It may often be a trivial and obvious matter to apply the sensitivity coefficients. For example, the temperature coefficient of linear expansion converts uncertainties in measured temperature with units of degrees into uncertainties with units of length. In the example just given the sensitivity coefficient for the measured length for temperature uncertainties is 11.5 ppm/°C for the steel tape or 1 ppm/°C for the invar tape.

The Drongo's Guide to Determining Measurement Uncertainty

Evaluation of the sensitivity coefficients can always be done by partial differentiation of the equation that models the measurement or by numerical calculation which approximates the differentiation process.

Consider an example where the value of a resistance, R_t , at a test temperature, t , is given by the equation:

$$R_t = R_0(1 + \alpha t) \quad (4)$$

where: α is the temperature coefficient of the resistor, ohms/°C, and

t is the temperature of the resistor in °C and

R_0 is the resistance in ohms at 0 °C.

We can differentiate the equation for R_t , with respect to t as follows:

$$\frac{\partial R_t}{\partial t} = R_0 \alpha \quad (5)$$

A similar process is used to determine all sensitivity coefficients.

It will be found to be convenient in some cases to calculate the measurand in terms of proportional parts (percent or parts per million) deviation from a nominal value. For example, a standard resistor might be found to be 73 ppm low compared to its nominal value. It would be appropriate to express the uncertainty in the same units, for example ± 2 ppm. Often, but not always, this procedure allows the use of the factor 1 as the sensitivity coefficients. Check the model equation to verify this.

It is often best to use the base SI units for all inputs as errors of three orders of magnitude are easily made when calculating sensitivity coefficients. A “reality check” should be made when doing an estimate for the first time. For example, the numeric value could be compared with that expected from experience.

7. Combined uncertainty

Once the individual components and their associated sensitivity coefficients have been evaluated, they may be combined to produce the combined standard uncertainty.

This is done by using equation (10) of the GUM. This equation must be modified if any of the components are correlated or there are significant higher order effects. These conditions are dealt with in the GUM.

The ISO GUM equation (10) can be written as:

$$u_c(y) = \sqrt{\sum (c_i u_i)^2} \quad (6)$$

where: $u_c(y)$ is the combined standard uncertainty of the measurand, and

The Drongo's Guide to Determining Measurement Uncertainty

c_i is the sensitivity coefficient for $u(x_i)$
which is the i th standard uncertainty of the input estimates and

Σ denotes the summation of all the terms, of which there are n .

This means that the uncertainty components are converted to the same units as the measurand using sensitivity coefficients then these products are squared and summed. The combined uncertainty is the square root of the sum. This is often called the “root sum of the squares” (RSS).

8. Expanded uncertainty

In order to have an adequate probability that the value of the measurand lies within the range given by the uncertainty the combined uncertainty is multiplied by a coverage factor. This coverage factor may be selected or it may be calculated so as to reflect a stated confidence level. For example, a factor of 2 gives an expanded uncertainty U with an approximate 95% confidence level.

This is acceptable, and is perhaps the best approach for testing situations where only a few worst case estimates are used for determining the test uncertainty. For high level calibrations the more rigorous method may be preferred.

When the uncertainty has been estimated from data with poor reliability, a larger coverage factor may be required to maintain a 95% confidence interval.

The expanded uncertainty is calculated from:

$$U_{95} = k u_c(y) \quad (7)$$

where: U_{95} is the expanded uncertainty at a
95% confidence limit and

k is the coverage factor, and

$u_c(y)$ is the combined standard uncertainty.

By assuming that the combined uncertainty has essentially a normal distribution we can use Student's t factor as the coverage factor k . This is justified by invoking the Central Limit Theorem, which in essence states that if many distributions are combined, irrespective of their own shape, the combined distribution will approximate a Normal distribution. Hence the combined uncertainty will tend towards a normal distribution as more and more components are included.

We can find values of k (= Student's t -factor) for any desired level of confidence provided we have the number of degrees of freedom. The question is how do we find the number of degrees of freedom for the combined standard uncertainty?

9. Effective degrees of freedom

Whilst the reason for determining the number of degrees of freedom associated with an uncertainty component is to allow the correct selection of the value of Student's t , it also gives an indication of how well a component may be relied upon. A high number of degrees of freedom is associated with a large number of measurements or a value with a low variance or low dispersion associated with it. A low number of degrees of freedom corresponds to a large dispersion or poorer confidence in the value.

Every component of uncertainty can have an appropriate number of degrees of freedom, ν , assigned to it.

For the mean, \bar{x} , for example, $\nu = n - 1$, where n is the number of repeated measurements.

For other Type A assessments, the process is also quite straightforward and uses established formula.

The obvious question is how to assign components evaluated by type B processes.

For some distributions, the limits may be determined so that we have complete confidence in their value. In such instances the number of degrees of freedom is effectively infinite. The assigning of limits which are worst case leads to this instance, namely infinite degrees of freedom, and simplifies the calculation of the effective degrees of freedom of the combined uncertainty.

If the limits themselves have some uncertainty, then a lesser number of degrees of freedom must be assigned. The ISO GUM gives a formula that is applicable to all distributions. It is equation G.3, reproduced below.

$$\nu \approx \frac{1}{2} \left(\frac{\Delta u(x)}{u(x_i)} \right)^{-2} \quad (8)$$

where: $\Delta u(x_i)$ is the relative uncertainty in the uncertainty and other terms as before.

Note that $\Delta u(x_i)$ is a number less than 1, but may for convenience be thought of as a percentage or a fraction. The smaller the number, the better defined is the magnitude of the uncertainty. For example, if the relative uncertainty is 10%, $\Delta u(x_i) = 0.1$ then it can be shown that the number of degrees of freedom, ν , is 50. For a relative uncertainty of 25% then $\nu = 8$ and for a relative uncertainty of 50%, ν is only 2.

Rather than become seduced by the elegance of the mathematics, it is better to try to determine the limits more definitely, particularly if the uncertainty is a major one.

It is of interest to note that equation (8) tells us that when we have made 51 measurements and taken the mean, the relative uncertainty in the uncertainty of the mean is 10%. This shows that even when many measurements are taken, the reliability of the uncertainty is not necessarily any better than when a type B assessment is made. Indeed, it is usually better to rely on prior knowledge rather than using an uncertainty based on two or three measurements. It also shows why we restrict the uncertainty to two digits. The value is usually not reliable enough to quote to better than 1% resolution.

The Drongo's † Guide to Determining Measurement Uncertainty

Once the uncertainty components have been combined, it remains to find the number of degrees of freedom in the combined uncertainty. The degrees of freedom for each component must also be combined to find the effective number of degrees of freedom to be associated with the combined uncertainty. This is calculated using the Welch-Satterthwaite equation, which is:

$$v_{eff} = \frac{u_c^4(y)}{\sum_{i=1}^n \frac{(c_i u_i)^4}{v_i}} \quad (8)$$

where: v_{eff} is the effective number of degrees of freedom for u_c , the combined uncertainty and

v_i is the number of degrees of freedom for u_i , the i th uncertainty term where

$u_i(y)$ is the product $c_i u(x_i)$, with the sign of c_i being neglected.

The other terms have their usual meaning.

11 Summary

- Make a model of the measurement system.
- List all the sources of uncertainties.
- Calculate the standard uncertainties for each component using type A analysis for those with repeated measurements and type B for others.
- Calculate the sensitivity coefficients.
- Calculate the combined uncertainty, and, if appropriate its effective degrees of freedom.
- Calculate the expanded uncertainty. Use a nominal or a calculated coverage factor. Round the measured value and the uncertainty to obtain the reported values.

† Drongo: Australian Slang for Idiot.