

Using repeated measurements to improve the standard uncertainty

Introduction

The standard uncertainty arising from random effects is often derived from repeated experiments and is quantified in terms of the standard deviation s of the measured values of the quantity. If it is the standard uncertainty on a single measurement of the quantity that is required then it is simply the observed standard deviation s ; but for a result which is the mean of n measurements, the standard uncertainty $u_{\bar{x}}$ will be reduced to the standard deviation of the mean:

$$u_{\bar{x}} = \frac{s}{\sqrt{n}} \quad \text{Eq. (1)}$$

Example 1 shows how Eq (1) applies when estimating the uncertainty of a mean value, but not when considering the uncertainty of individual observations.

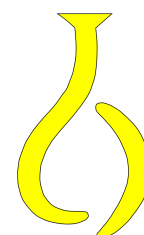
Example 1

A volumetric pipette is calibrated with 12 measurements; the mean and the standard deviation are calculated. Eq. 1 applies to estimate the standard uncertainty of the mean value. However, when *using* the pipette to deliver *one* aliquot, Eq. 1 does not apply and the standard uncertainty due to random variation on this single measurement is the standard deviation s

For Eq (1) to be valid, it is a requirement that all the measurements are independent and from a stable test sample under the **same** measurement conditions. The measurement conditions for all the measurements could be, for example, 1) repeatability conditions, 2) intermediate precision (within laboratory reproducibility) conditions or 3) reproducibility conditions

It is very important to be aware that the standard uncertainty given by Eq. 1 only gives the estimated uncertainty due to random variation under the conditions of measurement under which the observations were taken, and is only strictly valid for independent observations.

It can be hard to decide whether observations are sufficiently independent for Eq. 1 to apply, as there is no simple, general, rule. The following paragraphs therefore give examples of different circumstances to help identify circumstances when Eq. 1 can safely be used.



Eurachem

A FOCUS FOR
ANALYTICAL CHEMISTRY
IN EUROPE

Example where Equation (1) applies

Measurement of inhomogeneous samples

If inhomogeneity for test samples is a major part of the uncertainty the analyst may choose to measure more test portions from each test sample in order to reduce the standard uncertainty. If all these measurements are made under repeatability conditions, i.e. the same measurement conditions following the whole procedure, including repeated random sampling from the test sample, the standard deviation of the mean, given by Eq. (1), should be used to estimate the uncertainty arising from variation under repeatability conditions.

Examples where Equation (1) does not apply

The following paragraphs give two examples where neither the standard deviation nor the standard deviation of the mean can be directly applied without further analysis of the data.

Measurements taken in groups

One example is when we have data from internal quality control from a measurement procedure that includes a calibration before analysis each day. We want to calculate the standard uncertainty of the mean value used to set the central line in the QC chart. The data consist of duplicate QC measurements performed on a stable test sample each day over a longer period of time (say, p days), giving a total of $2p$ observations; that is, p groups of two measurements. Because each pair of duplicate measurements has a single calibration error in common; the duplicates within the data set are not strictly independent and Eq. (1) cannot be directly used for all $2p$ observations. The uncertainty of the mean can most easily be calculated by taking the standard deviation of the p means for each day and dividing that standard deviation by \sqrt{p} . Analysis of variance can also be useful in similar cases. Similar principles apply to other kinds of grouping, including grouping by operator, instrument etc.

Measurement when the test item or measurement system is not stable over time

Another common example is time dependent data. The time dependence could be due to instrument drift or to actual change in concentration by time. In such cases the error for a given observation is partly random and in part 'carried over' from the previous observation. Again the errors affecting each observation are not independent because some part of the error is common to successive observations; Eq. (1) can not be used and more complex statistical techniques, that allow for correlation, must be used to process the data.

For further guidance on the treatment of correlated data in uncertainty evaluation, see Eurolab Technical Report 1/2006: Guide to the Evaluation of Measurement Uncertainty for Quantitative Test Results, Appendix A.5 www.eurolab.org.