

Estimating and Combining Uncertainties

8th Annual ITEA Instrumentation Workshop
Park Plaza Hotel and Conference Center, Lancaster Ca.
5 May 2004

Howard Castrup, Ph.D.
President, Integrated Sciences Group
www.isgmax.com
hcastrup@isgmax.com

Abstract. A general measurement uncertainty model is presented that can be applied to measurements in which the value of an attribute is measured directly, and to multivariate measurements in which the value of an attribute is obtained by measuring various component quantities. In the course of the development of the model, axioms are developed that are useful to uncertainty analysis [1]. Measurement uncertainty is defined statistically and expressions are derived for estimating measurement uncertainty and combining uncertainties from several sources. In addition, guidelines are given for obtaining the degrees of freedom for both Type A and Type B uncertainty estimates.

Background

Suppose that we are making a measurement of some quantity such as mass, flow, voltage, length, etc., and that we represent our measured values by the variable x . Imagine that there is some "expectation value" for x , that we designate $\langle x \rangle$.¹ Each measured value x has some unknown error, that we designate ε_x . Given these designations, we can write each measured value as

$$x = \langle x \rangle + \varepsilon_x,$$

or

$$\varepsilon_x = x - \langle x \rangle.$$

The values we obtain for x may vary randomly from measurement to measurement due to several factors, such as environmental fluctuations, operator inconsistencies, etc. Because of their random nature,

¹ In this paper, the Dirac notation $\langle \cdot \rangle$ is used to indicate the expectation value of an observable variable or function.

If a variable x belongs to a population, its expectation value is synonymous with the population mean or average value. Thus, if x is a statistically distributed random variable, then $\langle x \rangle$ is the mean value of the distribution. In general, if $h(x)$ is a function of x , then $\langle h(x) \rangle$ is the value obtained by averaging h over all values of x in the population.

we treat these variations as statistical quantities and say that they follow statistical distributions.

We also acknowledge that the measurement reference we are using has some unknown bias. All we may know of this quantity is that it comes from some population of biases that also follow a statistical distribution. What knowledge we have of this distribution may derive from periodic calibration of the reference or from manufacturer data.

We are likewise cognizant of the fact that the finite resolution of our measuring system introduces another source of measurement error whose sign and magnitude are unknown and, therefore, can only be described statistically.

These observations lead us to a key axiom in uncertainty analysis:

Measurement errors are random variables that follow statistical distributions.

So, in measuring $\langle x \rangle$, we would say that x is distributed around $\langle x \rangle$. A graphical representation of this statement is shown in Figure 1. In the distribution shown, $\langle x \rangle = 10$. The function $f(x)$, called the *probability density function*, is related to the probability of occurrence of specific values of x . It is defined in such a way that, if δx is sufficiently small, the probability of obtaining a measured value between x and $x + \delta x$, is given by $f(x)\delta x$.²

In Figure 1, the spread of the distribution of measured values around the expectation value of 10 is quantified by the *standard deviation* σ_x of the distribution. For the case shown, $\sigma_x = 1$. If the standard deviation were smaller, the distribution would be less spread out. If it were larger, the spread would be greater.

² For the definition to be precise, δx needs to be infinitesimal.

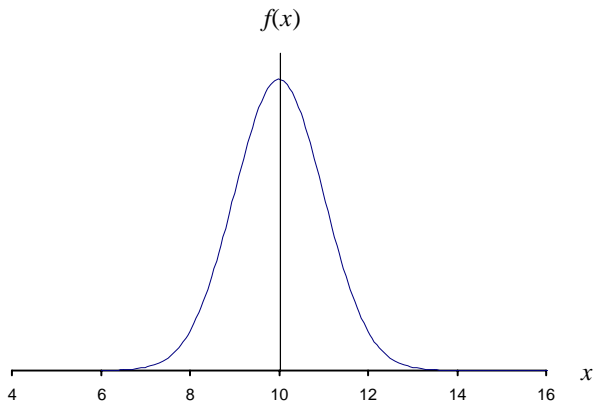


Figure 1. The distribution for measured values with an expectation value of 10 and a standard deviation or uncertainty of 1.

If the standard deviation were zero, the spread of the distribution would be zero, and we would say that all measured values x would be equal to the expectation value $\langle x \rangle$. For such measurements, we would say that we have zero measurement uncertainty.

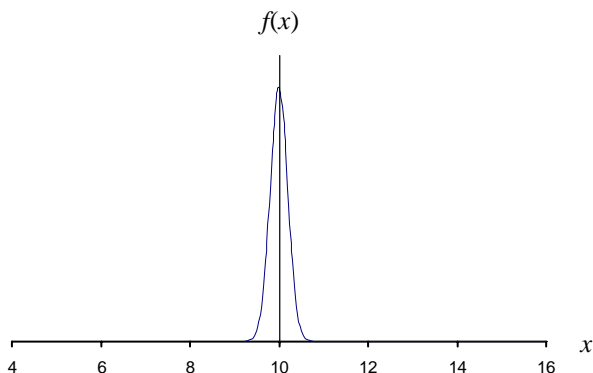


Figure 2. The distribution for measured values with an expectation value of 10 and a standard deviation or uncertainty of 0.2.

If the spread of the distribution is nonzero but small, as in Figure 2, we would obtain measured values that are confined to a small neighborhood around the expectation value. In this case, we would say that the measurement uncertainty is small but not zero.

We can readily see that if the distribution were widely spread, as in Figure 3, then the values for x that we might obtain by measurement would vary considerably. In this case, we would say that the measurement uncertainty is large.

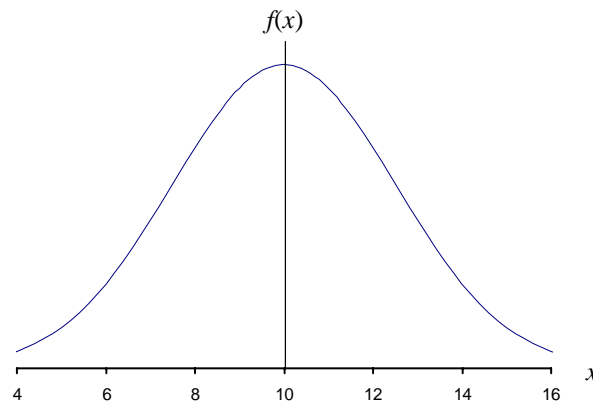


Figure 3. The distribution for measured values with an expectation value of 10 and a standard deviation or uncertainty of 2.5.

Defining Uncertainty

We have seen that the uncertainty in a measured value x is a measure of the extent that values of x are spread around the expectation value $\langle x \rangle$. Another way of saying this is that the more spread out the distribution around $\langle x \rangle$, the larger the uncertainty. Since the deviations of x from $\langle x \rangle$ are just the measurement errors, given by $\epsilon_x = x - \langle x \rangle$, we expect that the uncertainty is in some way related to the measurement errors ϵ_x .

We might at first suppose that the uncertainty is some kind of average value for ϵ_x . If so, we would write

$$u_x \sim \text{average}(x - \langle x \rangle) = \langle x - \langle x \rangle \rangle,$$

This is a conceptually palatable definition that, unfortunately, does not work. It turns out that this quantity is always zero:

$$\langle x - \langle x \rangle \rangle = \langle x \rangle - \langle x \rangle = 0$$

So, we need another definition. The first one didn't work because the "negatives" cancel the "positives." Clearly, we need an unsigned quantity. One candidate might be the square of the deviations:

$$u_x \sim \text{average}[(x - \langle x \rangle)^2] = \langle (x - \langle x \rangle)^2 \rangle.$$

This quantity is the average or mean value of the error squared. Accordingly, it is called the *mean square error*.

Since $(x - \langle x \rangle)^2$ is always positive, its average value does *not* always equal zero. In addition, we can easily see that, if a distribution is widely spread, the mean square error will be large. Conversely, if a distribution is tightly grouped, the mean square error will be small. We appear to be on the right track. The only problem is that the mean square error is not in the same units as the measured values. For instance if we are measuring

volts, the mean square error is expressed in volts-squared.

So, while the mean square error possesses some of the properties we seek in a definition of measurement uncertainty, it is not in itself an acceptable definition. But how about the square root of the mean square error? This quantity has the properties we want and is expressed in the same units as measured values. In fact, all the requirements for a good definition of uncertainty are met by the square root, and we write

$$u_x = \sqrt{\langle (x - \langle x \rangle)^2 \rangle}. \quad (1)$$

We note that, by definition, the square root of the mean square error is just the standard deviation. Accordingly,

$$u_x = \sigma_x,$$

and we have a second key axiom of uncertainty analysis:

The uncertainty in the measurement of $\langle x \rangle$ is equal to the standard deviation of the distribution of measurements of $\langle x \rangle$.

Substituting $\varepsilon_x = x - \langle x \rangle$ in Eq. (1) gives

$$u_x = \sqrt{\langle \varepsilon_x^2 \rangle}.$$

Since the square root of the mean square error in a quantity is the standard deviation for the quantity, then, by the foregoing axiom, we can write

$$\begin{aligned} u_x &= \sqrt{\langle \varepsilon_x^2 \rangle} \\ &= \sigma_{\varepsilon_x} \\ &= u_{\varepsilon_x}. \end{aligned} \quad (2)$$

This expression provides a third (and final) key axiom of uncertainty analysis:

The uncertainty in the measurement of $\langle x \rangle$ is equal to the uncertainty in the measurement error.

Combining Uncertainties

Suppose, for discussion purposes, that the error in x is composed of errors from two sources: measurement parameter bias and measurement process random (repeatability) error.³ We then have

³ This is not to suggest that these error sources are the only ones present in all measurements. There may also be errors due to resolution, environment, operator, etc.

$$x = \langle x \rangle + \varepsilon_{xb} + \varepsilon_{xr},$$

where ε_{xb} is the error due to parameter bias and ε_{xr} is the random error. Given the above definition of measurement uncertainty and invoking the third axiom, we have

$$\begin{aligned} u_x^2 &= \langle (\varepsilon_{xb} + \varepsilon_{xr})^2 \rangle \\ &= \langle \varepsilon_{xb}^2 + \varepsilon_{xr}^2 + 2\varepsilon_{xb}\varepsilon_{xr} \rangle \\ &= \langle \varepsilon_{xb}^2 \rangle + \langle \varepsilon_{xr}^2 \rangle + 2\langle \varepsilon_{xb}\varepsilon_{xr} \rangle. \end{aligned}$$

Applying Eq. (2), we can write the above expression as

$$u_x^2 = u_{xb}^2 + u_{xr}^2 + 2\langle \varepsilon_{xb}\varepsilon_{xr} \rangle.$$

So, we see that the square of the total uncertainty due to measurement error is equal to the sum of the squares of the uncertainties of the constituent errors plus an additional term. The additional term is called the *covariance* of ε_{xb} and ε_{xr} .

To simplify the analysis, we usually define another variable, called the *correlation coefficient* ρ_{br} as

$$\rho_{br} \equiv \frac{\langle \varepsilon_{xb}\varepsilon_{xr} \rangle}{u_{xb}u_{xr}}.$$

With this definition, we can now write

$$u_x^2 = u_{xb}^2 + u_{xr}^2 + 2\rho_{br}u_{xb}u_{xr},$$

and

$$u_x = \sqrt{u_{xb}^2 + u_{xr}^2 + 2\rho_{br}u_{xb}u_{xr}}. \quad (3)$$

Statistically Independent Errors

For most measurements, the bias of the measuring parameter and the error due to random variations are completely independent of one another. If so, it can be shown that

$$\langle \varepsilon_{xb}\varepsilon_{xr} \rangle = 0,$$

and, consequently, that $\rho_{br} = 0$. In this case, we have

$$u_x = \sqrt{u_{xb}^2 + u_{xr}^2}, \quad (4)$$

i.e., the total uncertainty in the measurement of x is the square root of the sum of the constituent uncertainties. The latter quantity is called the *root sum square* or *rss* of the constituent uncertainties.

In general, we can make the following assertion:

The total uncertainty in the sum of independent errors is a root-sum-square (rss) combination of their uncertainties.

Highly Correlated Errors

In making complex measurements involving more than one variable, we sometimes encounter cases where the correlation coefficient is not zero. Suppose, for instance that we are interested in the uncertainty in a quantity

$$z = x + y,$$

obtained by direct measurement of the quantities x and y . We first write

$$\begin{aligned} z &= \langle x \rangle + \varepsilon_x + \langle y \rangle + \varepsilon_y \\ &= \langle x \rangle + \langle y \rangle + \varepsilon_x + \varepsilon_y \\ &= \langle z \rangle + \varepsilon_x + \varepsilon_y, \end{aligned}$$

so that

$$\varepsilon_z = \varepsilon_x + \varepsilon_y$$

and

$$\begin{aligned} u_z^2 &= \langle \varepsilon_z^2 \rangle = \langle (\varepsilon_x + \varepsilon_y)^2 \rangle \\ &= \langle \varepsilon_x^2 \rangle + \langle \varepsilon_y^2 \rangle + 2\langle \varepsilon_x \varepsilon_y \rangle \\ &= u_x^2 + u_y^2 + 2\rho_{xy}u_xu_y. \end{aligned}$$

If both x and y are so strongly correlated that $\rho_{xy} = 1$. Then the expression for u_z becomes

$$\begin{aligned} u_z &= \sqrt{u_x^2 + u_y^2 + 2u_xu_y} \\ &= \sqrt{(u_x + u_y)^2} \\ &= u_x + u_y. \end{aligned}$$

From this result, we can make the following statement:

If the correlation coefficient between errors is equal to one, the total uncertainty of their sum is the sum of the individual uncertainties.

Note that, if the correlation coefficient is equal to minus one, we have

$$\begin{aligned} u_z &= \sqrt{u_x^2 + u_y^2 - 2u_xu_y} \\ &= \sqrt{(u_x - u_y)^2} \\ &= |u_x - u_y|, \rho_{xy} = -1. \end{aligned}$$

Multivariate Measurements and Sensitivity Coefficients

The general error model for measurements involving more than one variable expresses the total error as a weighted sum of the constituent errors. To develop this model, we again consider the measurement of a quantity z obtained by measuring two quantities x and y . In this case, however, x and y are related to z by some function

$$z = h(x, y).$$

The general error model for the total bias ε_z is written

$$\begin{aligned} \varepsilon_z &= \left(\frac{\partial h}{\partial x}\right)\varepsilon_x + \left(\frac{\partial h}{\partial y}\right)\varepsilon_y, \\ &= c_x\varepsilon_x + c_y\varepsilon_y, \end{aligned}$$

where c_x and c_y are referred to as *sensitivity coefficients*. Extrapolating from the previous section, we can write the total uncertainty as

$$\begin{aligned} u_z &= \sqrt{\langle (c_x\varepsilon_x + c_y\varepsilon_y)^2 \rangle} \\ &= \sqrt{\langle c_x^2\varepsilon_x^2 + c_y^2\varepsilon_y^2 + 2c_xc_y\varepsilon_x\varepsilon_y \rangle} \\ &= \sqrt{c_x^2\langle \varepsilon_x^2 \rangle + c_y^2\langle \varepsilon_y^2 \rangle + 2c_xc_y\langle \varepsilon_x\varepsilon_y \rangle} \\ &= \sqrt{c_x^2u_x^2 + c_y^2u_y^2 + 2c_xc_y\rho_{xy}u_xu_y}. \end{aligned}$$

Again, if the x and y measurements are independent of one another, then $\rho_{xy} = 0$, and the uncertainty becomes

$$u_z = \sqrt{c_x^2u_x^2 + c_y^2u_y^2}.$$

On the other hand, if the biases in the x and y measurements are correlated such that $\rho_{xy} = 1$, then the uncertainty becomes

$$\begin{aligned} u_z &= \sqrt{c_x^2u_x^2 + c_y^2u_y^2 + 2c_xc_yu_xu_y} \\ &= \sqrt{(c_xu_x + c_yu_y)^2} \\ &= c_xu_x + c_yu_y. \end{aligned}$$

Of course, this treatment of uncertainty can be extended to cases involving any number of errors and uncertainties.

Combining Type A and Type B Estimates

An uncertainty estimate computed as the standard deviation of a random sample of measurements or determined by analysis of variance is called a *Type A estimate*. An uncertainty estimate determined

heuristically, in the absence of sampled data, is called a *Type B estimate*.

The current mindset is that a Type A estimate is a “statistical” quantity, whereas a Type B estimate is not. The main reason for this is that we can quantify a Type A estimate by the amount of information that went into calculating it, whereas it is commonly believed that we can’t do the same for a Type B estimate. The amount of information used to estimate the uncertainty in a given error is called the *degrees of freedom*. The degrees of freedom is required, among other things, to employ an uncertainty estimate in computing confidence limits commensurate with some desired confidence level.

For a Type A estimate, the degrees of freedom is the size of the random sample minus one. Since a Type B estimate is, by definition, obtained without recourse to a sample of data, we obviously don’t have a sample size to work with. However, we *can* develop something analogous to a sample size by applying a method based on Eq. G.3 of the GUM [2].

This method involves extracting what is known about a given measurement error and then converting this information into an effective degrees of freedom [3-5].⁴

Component Uncertainty

In multivariate measurements, the variables that are measured to obtain a value for a quantity of interest are referred to as *component variables*, and the errors in their measurements are called *component errors*. In estimating the uncertainty in the measurement of each component error, we break the error down into constituent *error sources* and estimate the uncertainty due to each error source.

A direct measurement is treated in the same way, except that, in a direct measurement, there is only one component error.

Suppose that we have a component error ε_c whose error sources include measuring parameter bias, random error, resolution error and operator bias. The error model for this measurement can be expressed as

$$\varepsilon_c = \varepsilon_{bias} + \varepsilon_{random} + \varepsilon_{resolution} + \varepsilon_{operator} .$$

If we apply the uncertainty estimation expressions developed in this paper, we have

⁴ The method assumes, as with most statistical tools in the GUM and elsewhere, that measurement errors are approximately normally distributed.

$$u_c^2 = u_{bias}^2 + u_{random}^2 + u_{resolution}^2 + u_{operator}^2 ,$$

where it can be safely assumed that the various errors are independent of one another, i.e., the correlation coefficients are all zero. From this expression, we can clearly see that uncertainties are combined without regard for how they are estimated, i.e., without distinguishing between Type A and Type B estimates. However, after we have obtained estimates of each of the constituent uncertainties, does the total estimate mean anything? Can it be *used* in any way, such as to establish confidence limits, for example?

As indicated in the opening remarks in this section, unless we can obtain the degrees of freedom for each constituent, the answer is no — at least as far as developing confidence limits is concerned. It therefore behooves us to determine the degrees of freedom for each uncertainty estimate.

Random Error

For u_{random} , this is easy. From earlier, we observed that the uncertainty in a measurement is equal to the standard deviation of the measurement error distribution. We can approximate this standard deviation by taking a random sample of measurements and computing the *sample* standard deviation. The appropriate expression is

$$u_{random} \approx s_x = \sqrt{\frac{1}{n-1} \sum_{i=1}^n (x_i - \bar{x})^2} ,$$

where x_i is the *i*th measured value in the sample, n is the sample size, \bar{x} is the sample mean and s_x is the sample standard deviation. The “approximately equal” sign in this expression indicates that the measurement sample is finite. That is, that the amount of information that went into estimating the uncertainty due to random error is not infinite. This amount of information is the degrees of freedom ν_{random} . For the above estimate, it is equal to $n - 1$:

$$\nu_{random} = n - 1 .$$

Resolution Error

For $u_{resolution}$, we have two possibilities; we either have analog readings or a digital display.

Digital Display

If we have a digital display, errors due to resolution are contained within \pm half the resolution increment around the indicated value. In addition, the probability of the resolution being anywhere within these limits is equal. That is, there is no special likelihood that a resolution

error will be near the limits or near the indicated value or anywhere else within the limits. The statistical distribution for such errors is the rectangular or *uniform* distribution shown below.⁵

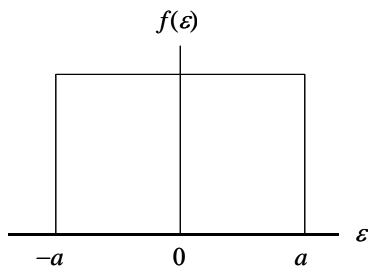


Figure 4. The Uniform or Rectangular Error Distribution: The probability of lying between the distribution limits $-a$ and a is constant. The probability of lying outside $\pm a$ is zero.

The standard deviation or uncertainty for the uniform distribution is equal to the distribution limit a divided by the square root of three. Hence, the uncertainty due to resolution error is just half the last digit divided by the root three. For example, imagine that a digital readout is to three decimal places. Then the uncertainty due to resolution error is

$$u_{resolution} = \frac{a}{\sqrt{3}} = \frac{0.0005}{\sqrt{3}} = 0.00029.$$

Note that there is no ambiguity in obtaining this number. The “amount of information” is all there is on the subject. Accordingly, for this error source, the degrees of freedom for the uncertainty estimate is infinite.

Analog Display

If we have an analog display, some judgment is needed by the operator in estimating limits $\pm L$ within which readings can be resolved. The operator must also estimate a probability or confidence level p that resolution errors are contained within these limits. In such cases, we assume that resolution errors are normally distributed. If we are certain of the estimated limits and the confidence level, then the resolution uncertainty is computed using

$$u_{resolution} = \frac{L}{\Phi^{-1}\left(\frac{1+p}{2}\right)},$$

where the operator Φ^{-1} is the inverse normal distribution function.

⁵ For a discussion on error distributions, see [6].

When we are certain of the estimated limits and the confidence level, the degrees of freedom are infinite. If the limits or the confidence level are not precisely known, we must apply what is known about the resolution error to estimating an effective degrees of freedom [3 – 5]. For example, suppose the operator claims that $L = 50\%$ of the smallest increment of resolution (give or take 10%) and that he can resolve readings within $\pm L$ nine times out of ten. If so, the effective degrees of freedom are⁶

$$v_{resolution} \approx 38.$$

Measurement Bias

We typically assume that u_{bias} , is normally distributed about the measuring parameter's reading or its nominal value, whichever applies. With regard to the degrees of freedom in u_{bias} , suppose that we have some knowledge of the calibration history for the parameter and can make the following statement: “In calibrating the parameter used to measure the quantity of interest, we have found it to be in-tolerance 90% of the time.” If we are using the measuring parameter near the end of its calibration interval, and the 90% figure is based on a large number of prior calibrations, then there is little ambiguity in our knowledge, and the degrees of freedom for u_{bias} can be taken to be essentially infinite. But if there is some “fuzziness” in the 90% number, due to the sparseness of calibration history or due to using the parameter at some intermediate point in its interval, this would not be so — the degrees of freedom would be finite. For instance, let’s say that, for this particular measurement parameter, we estimate the in-tolerance figure to be 90%, give or take 5%. If so, then we have

$$v_{bias} \approx 69.$$

Operator Bias

For $u_{operators}$, as with u_{bias} , we rely on prior history. In this case, the prior history may consist of our recollected experience in making the measurement of interest. Since we are dealing with operator bias, which is expressed as a tendency to slant measurement results one way or another, we may be able to formulate a set of bounding limits such that the errors due to operator tendencies can be said to lie within the limits with some definable containment probability. For example, imagine that we believe that operator bias is confined to \pm limits of between 8 and 10, with a containment

⁶ All Type B degrees of freedom estimates presented in this paper were obtained using the ISG Type B Uncertainty Calculator [4].

probability of 75%, give or take 10%. Then the degrees of freedom due to operator bias is⁷

$$v_{operator} \approx 26.$$

Component Degrees of Freedom

We now have degrees of freedom for both Type A and Type B estimates. This allows us to treat the total component uncertainty as a statistical quantity. The degrees of freedom for the total uncertainty is obtained using the Welch-Satterthwaite relation [2]. In general, if a component error is composed of k error sources ε_i , with uncertainties u_i , $i = 1, 2, \dots, k$, this relation gives the degrees of freedom v_c as

$$v_c \approx \frac{u_c^4}{\sum_{i=1}^k \frac{u_i^4}{v_i}}.$$

For the present discussion, this becomes

$$v_c \approx \frac{(u_{bias}^2 + u_{random}^2 + u_{resolution}^2 + u_{operator}^2)^2}{\frac{u_{bias}^4}{69} + \frac{u_{random}^4}{n-1} + \frac{u_{resolution}^4}{\infty} + \frac{u_{operator}^4}{26}}.$$

Conclusion

It was asserted that measurement errors are random variables that are statistically distributed around expectation values. This assertion was expressed as the axiom *Measurement errors are random variables that follow statistical distributions.*

A property of error distributions that is important to uncertainty analysis is the distribution's *standard deviation*. The standard deviation quantifies the spread of errors around the expectation value. Because of its characteristics, e.g., the wider the spread, the larger the standard deviation, it is the ideal property to use for measurement uncertainty. This was also stated as an axiom: *The uncertainty in the measurement of $\langle x \rangle$ is equal to the standard deviation of the distribution of measurements of $\langle x \rangle$.*

⁷ The ISG Type B Uncertainty Calculator was again employed. It should be mentioned that parameter bias uncertainty, random uncertainty and operator bias can all be calculated as Type A estimates using analysis of variance with sampled data. However, with regard to operator bias, the number of different operators that are sampled is usually so small as to invalidate the operator bias uncertainty estimate. In such cases, you are often better off attempting to make a Type B estimate.

Acceptance of these axioms and the application of Eqs. (1) and (2), produced a third axiom: *The uncertainty in a measurement is equal to the uncertainty in the measurement error.*

This axiom is especially useful in uncertainty analysis, particularly in accounting for correlations. When considering correlations, it is important to bear in mind that the relevant correlations are those between measurement errors rather than measured quantities. Thus, two quantities may be functionally related and, seemingly, correlated but their measurement errors may be completely independent and uncorrelated.

Working within the framework of the above axioms, a methodology was presented that lead to expressions for both estimating measurement uncertainty and combining uncertainties from several sources. The methodology is applicable to both direct and multivariate measurements.

It was argued that, to use an uncertainty estimate in a statistical manner, its degrees of freedom must be determined. For a Type A estimate, based on a measurement sample, the degrees of freedom value is the sample sizes minus one. The degrees of freedom value for a Type B estimate, is obtained using a method developed by Integrated Sciences Group [3] and implemented in shareware [4].

References

- [1] Castrup, S., "A Comprehensive Comparison of Uncertainty Analysis Tools," *Proc. Measurement Science Conference*, Anaheim, January 2004.
- [2] ANSI/NCSL Z540-2-1997, *U.S. Guide to the Expression of Uncertainty in Measurement*, NCSL, Boulder, 1997.
- [3] Castrup, H., "Estimating Category B Degrees of Freedom," *Proc. Measurement Science Conference*, Anaheim, January 2000.
- [4] ISG Type B Uncertainty Calculator, © 1999-2003, Integrated Sciences Group. All rights reserved.
- [5] UncertaintyAnalyzer, © 1994-2004, Integrated Sciences Group, all rights reserved.
- [6] Castrup, H., "Distributions for Uncertainty Analysis," *Proc. International Dimensional Workshop*, Knoxville, May 2001.