Comparison of Methods for Establishing Confidence Limits and Expanded Uncertainties

Suzanne Castrup
Integrated Sciences Group
14608 Casitas Canyon Rd.
Bakersfield, CA 93306
661-872-1683
scastrup@isgmax.com

***Abstract:** In reporting measurement results, it may be necessary to include an interval that contains the true value with some specified confidence level or probability. The interval may be reported as confidence limits, with an associated confidence level, or an expanded uncertainty, with an associated coverage factor. This paper examines three methods for computing confidence limits and expanded uncertainties: 1) GUM, 2) Convolution and 3) Monte Carlo Simulation. The first method combines error distribution variances, while the second and third methods directly combine error distributions via mathematical or numerical techniques. Four direct measurement scenarios are evaluated and the intervals computed from each method are compared.*

INTRODUCTION

Measurement uncertainty plays an important role in making decisions, managing risk, developing tolerances, selecting measurement methods, developing capability statements, achieving laboratory accreditation, hypothesis testing, establishing calibration intervals and communicating technical variables. Therefore, uncertainty estimates should realistically reflect the measurement process under investigation or evaluation.

Measurement uncertainty must also be reported in a way that can be readily understood and interpreted by others. At a minimum, the measured value, the combined standard uncertainty, its estimate type (A, B or A/B) and degrees of freedom should be reported. In some instances, confidence limits with an associated confidence level or an expanded uncertainty with associated coverage factor may also be reported.

Three methods for computing confidence limits and expanded uncertainties are discussed and compared. The GUM[1] method evaluates and combines error distribution variances and calculates the effective degrees of freedom of the uncertainty estimate for the combined measurement error. The convolution method uses a mathematical approach for combining error distributions. The Monte Carlo method involves the combination of error distributions via numerical simulation. All three methods require the identification of measurement process errors and the evaluation of error distributions. Consequently, appropriate error distributions must be applied to achieve realistic results from any of these methods.

UNCERTAINTY ESTIMATION

A measurement is a process whereby the value of a quantity is estimated. In any given measurement scenario, each measured quantity is also accompanied by measurement error. The basic relationship between the measured quantity x and the measurement error ε_x is given in equation (1).

$$x = x_{\text{true}} + \varepsilon_x \tag{1}$$

An error model is an algebraic expression that defines the total error in the value of a measured quantity in terms of all relevant measurement process errors. The error model for ε_x is the sum of the errors encountered during the measurement of x

[1] Through out this paper, the term GUM refers to ISO *Guide to the Expression of Uncertainty in Measurement* and ANSI/NCSLI Z540-2-1997, the American National Standard for Expressing Uncertainty – U.S. *Guide to the Expression of Uncertainty in Measurement*.

$$\varepsilon_x = \varepsilon_1 + \varepsilon_2 + \cdots + \varepsilon_k \qquad (2)$$

where the numbered subscripts signify the different measurement process errors. Measurement uncertainty is a quantification of our lack of knowledge of the sign and magnitude of measurement error.

Measurement Process Errors

Measurement process errors are the basic elements of uncertainty analysis. Once these fundamental error sources have been identified, we can begin to develop uncertainty estimates. The errors most often encountered in making measurements include, but are not limited to the following:

- Reference Attribute Bias
- Repeatability
- Resolution Error
- Computation Error
- Operator Bias
- Environmental Factors Error

Reference Attribute Bias

Calibrations are performed to obtain an estimate of the value or bias of selected unit-under-test (UUT) attributes by comparison to corresponding measurement reference attributes. The error in the value of a reference attribute, at any instant in time, is composed of a systematic component and a random component. Reference attribute bias is the systematic error component that persists from measurement to measurement during a measurement session.[2] Attribute bias excludes resolution error, random error, operator bias and other error sources that are not properties of the attribute.

Repeatability

Repeatability is a random error that manifests itself as differences in measured value from measurement to measurement during a measurement session. It is important to note that, random variations in a measured quantity or UUT attribute are not separable from random variations in the reference attribute or random variations due to other error sources.

Resolution Error

Reference attributes and/or UUT attributes may provide indications of sensed or stimulated values with some finite precision. The smallest discernible value indicated in a measurement comprises the resolution of the measurement. For example, a voltmeter may indicate values to four, five or six significant digits. A tape measure may provide length indications in meters, centimeters or millimeters. A scale may indicate weight in terms of kg, g, mg or µg.

The basic error model for resolution error, ε_{res}, is

$$\varepsilon_{res} = x_{indicated} - x_{sensed}$$

where x_{sensed} is a "measured" value detected by a sensor or provided by a stimulus and $x_{indicated}$ is the indicated representation of x_{sensed}. In some measurement situations, repeatability may be considered to be a manifestation of resolution error. If measurement repeatability is smaller than the display resolution, only resolution error should be included in the uncertainty analysis. If measurement repeatability is larger than the display resolution, then both error sources should be included in the uncertainty analysis.

Operator Bias

Errors can be introduced by the person or operator making the measurement. Because of the potential for human operators to acquire measurement information from an individual perspective or to produce a systematic bias in a measurement result, it sometimes happens that two operators observing the same measurement result will

[2] A measurement session is considered to be an activity in which a measurement or sample of measurements is taken under fixed conditions, usually for a period of time measured in seconds, minutes or, at most, hours.

systematically perceive or produce different measured values.

In reality, operator bias has a somewhat random character due to inconsistencies in human behavior and response. The random contribution is included in measurement repeatability and the systematic contribution is the operator bias.

Environmental Factors Error

Errors can result from variations in environmental conditions, such as temperature, vibration, humidity or stray emf. Additional errors are introduced when measurement results are corrected for environmental conditions. For example, when correcting a length measurement for thermal expansion, the error in the temperature measurement will introduce an error in the length correction. The uncertainty in the correction error is a function of the uncertainty in the error in the environmental factor.[3]

Computation Error

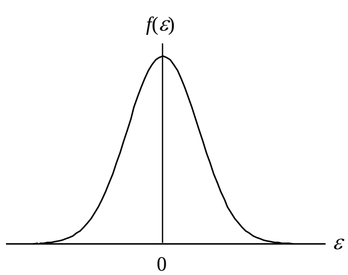
Data processing errors result from computation round-off or truncation, numerical interpolation of tabulated values, or the use of curve fit equations. For example, in the regression analysis of a range of values, the standard error of estimate quantifies the difference between the measured values and the values estimated from the regression equation.[4]

A regression analysis that has a small standard error of estimate has data points that are very close to the regression line. Conversely, a large standard error of estimate results when data points are widely dispersed around the regression line. However, if another sample of data were collected, then a different regression line would result. The standard error of the forecast accounts for the dispersion of various regression lines that would be generated from multiple sample sets around the true population regression line. The standard error of forecast is a function of the standard error of estimate and the measured value and should be used when estimating uncertainty due to regression error.

Error Distributions

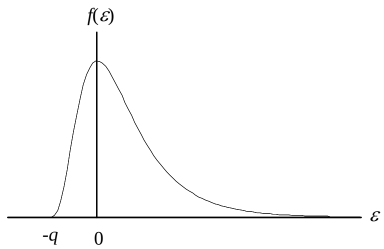
An important aspect of the uncertainty analysis is the fact that measurement errors can be characterized by probability distributions. The probability distribution for a type of measurement process error is a mathematical description of how likely an error or a range of errors is likely or unlikely to occur. With a basic understanding of error distributions and their statistics, we can estimate uncertainties. Error distributions include, but are not limited to normal, lognormal, uniform (rectangular), triangular, quadratic, cosine, exponential, U-shaped and Student's t. Probability density functions for selected distributions are summarized in Table 1.

Table 1. Probability Distributions for Measurement Process Errors

Probability Distribution	**Distribution Plot**	**Probability Density Function and Uncertainty Equation**
Normal		$f(\varepsilon) = \frac{1}{\sqrt{2\pi}\sigma} e^{-(\varepsilon)^2 / 2\sigma^2}$ $u = \frac{L}{\Phi^{-1}\left(\frac{1+p}{2}\right)}$ where σ is the standard deviation, $\pm L$ are the containment limits, p is the containment probability and $\Phi^{-1}()$ is the inverse normal distribution function.

[3] In the length correction scenario, error in the coefficient of thermal expansion may also need to be taken into account.

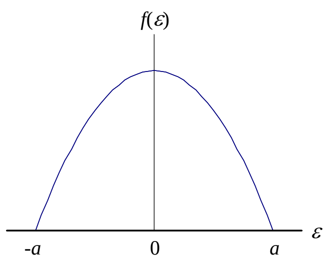
[4] Hanke, J. et al.: *Statistical Decision Models for Management*, Allyn and Bacon, Inc. 1984.

<table>
<tr><th>Probability Distribution</th><th>Distribution Plot</th><th>Probability Density Function and Uncertainty Equation</th></tr>
<tr><td>Lognormal</td><td>


</td><td>

$$f(\varepsilon) = \frac{1}{\sqrt{2\pi}\lambda|\varepsilon + q|}\exp\left\{-\frac{\left[\ln\left(\frac{\varepsilon + q}{m + q}\right)\right]^2}{2\lambda^2}\right\}$$

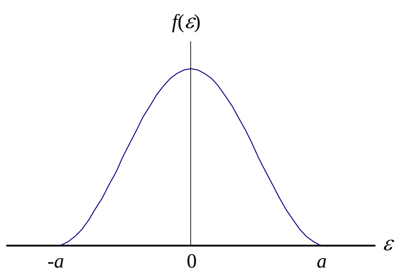
$$u_\varepsilon = |m + q| e^{\lambda^2/2}\sqrt{e^{\lambda^2} - 1}$$

where q is a physical limit for ε, m is the distribution median, and λ is a shape parameter.
</td></tr>
<tr><td>Quadratic</td><td>


</td><td>

$$f(\varepsilon) = \begin{cases} \frac{3}{4a}\left[1 - (\varepsilon / a)^2\right], & -a \le \varepsilon \le a \\ 0, & \text{otherwise} \end{cases}$$

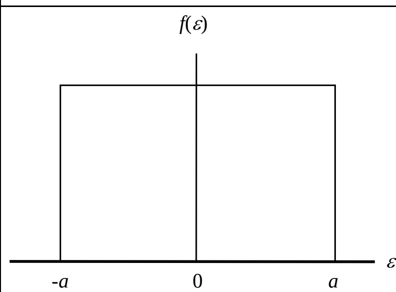
$$u_\varepsilon = \frac{a}{\sqrt{5}}$$

where ± a are the minimum distribution bounding limits.
</td></tr>
<tr><td>Cosine</td><td>


</td><td>

$$f(\varepsilon) = \begin{cases} \frac{1}{2a}\left[1 + \cos(\pi\varepsilon / a)\right], & -a \le \varepsilon \le a \\ 0, & \text{otherwise} \end{cases}$$

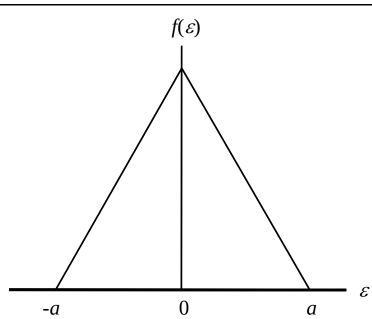
$$u_\varepsilon = \frac{a}{\sqrt{3}}\sqrt{1 - \frac{6}{\pi^2}}$$

where ± a are the minimum distribution bounding limits.
</td></tr>
<tr><td>Uniform</td><td>


</td><td>

$$f(\varepsilon) = \begin{cases} \frac{1}{2a}, & -a \le \varepsilon \le a \\ 0, & \text{otherwise} \end{cases}$$

$$u_\varepsilon = \frac{a}{\sqrt{3}}$$

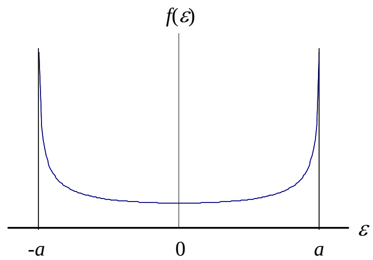
where ± a are the minimum distribution bounding limits.
</td></tr>
<tr><td>Triangular</td><td>


</td><td>

$$f(\varepsilon) = \begin{cases} (\varepsilon + a)/a^2, & -a \le \varepsilon \le 0 \\ (a - \varepsilon)/a^2, & 0 \le \varepsilon \le a \\ 0, & \text{otherwise} \end{cases}$$

$$u_\varepsilon = \frac{a}{\sqrt{6}}$$

where ± a are the minimum distribution bounding limits.
</td></tr>
</table>

Probability Distribution	Distribution Plot	Probability Density Function and Uncertainty Equation
U-Shaped		$f(\varepsilon)=\begin{cases}\dfrac{1}{\pi\sqrt{a^2+\varepsilon^2}}, & -a\le\varepsilon\le a\\ 0, & \text{otherwise}\end{cases}$ $u_\varepsilon=\dfrac{a}{\sqrt{2}}$ where ± *a* are the minimum distribution bounding limits.

Choosing the Appropriate Distribution

To obtain realistic uncertainty estimates, it is important that appropriate error distributions are used. The normal and lognormal distributions are relevant to most real world measurement applications. Other distributions such as the uniform, triangular, quadratic, cosine and U-shaped are also possible, although they have limited applicability.

The normal distribution applies to a wide variety of measurement process errors, and is often used as the default distribution, unless information to the contrary is available. The lognormal distribution is applied for errors that are bounded on one side with skewed (i.e., non-symmetric) containment limits. When using the normal or lognormal distribution, some effort must be made to establish containment limits and containment probability.

The uniform distribution is applicable for only a few measurement processes errors. Chief among these are resolution error resulting from a digital readout or quantization error due to analog to digital conversion. This results from the fact that it is hard to find real world instances in which a measurement process error has an equal probability of occurrence between two limits and zero probability of occurrence outside of these limits. When applying the uniform distribution, it is important that the 100% containment limits are known and that they represent the minimum bounding limits.

The quadratic, cosine or triangular distributions are applicable for errors that exhibit a central tendency, have 100% containment and known minimum bounding limits. The U-shaped distribution is applicable for errors that have 100% containment limits and the highest probability of occurrence is at or near known minimum bounding limits.

Measurement Process Uncertainty

Measurement processes uncertainty is the standard deviation of the probability distribution for the measurement process error. The standard deviation of an error distribution is the square root of the distribution's *statistical variance*. If $f(\varepsilon)$ represents the probability density function for an error distribution, and μ_ε represents the distribution's mode or mean value, then the variance or mean square error is given by

$$\text{var}(\varepsilon)=\int_{-\infty}^{\infty} f(\varepsilon)(\varepsilon-\mu_\varepsilon)^2\,d\varepsilon \qquad (3)$$

and the standard deviation becomes

$$u_\varepsilon=\sqrt{\text{var}(\varepsilon)}\,. \qquad (4)$$

For symmetric error distributions, μ_ε is taken to be zero. In these cases, equation (3) reduces to

$$\text{var}(\varepsilon)=\int_{-\infty}^{\infty} f(\varepsilon)\,\varepsilon^2\,d\varepsilon \qquad (5)$$

Equations (4) and (5) show that the uncertainty in a measurement error can be estimated if the distribution's probability density function is known.

There are two approaches to estimating the standard deviation of an error distribution. Type A estimates involve data sampling and analysis. Type B estimates use engineering knowledge or recollected experience of measurement processes.

Type A Estimates

Random sampling is a cornerstone for obtaining relevant statistical information for Type A uncertainty estimates. Therefore, Type A estimates usually apply to the uncertainty due to repeatability or random error. The data used for Type A uncertainty estimates typically consist of sampled values.[5] It is important that each repeat measurement is independent, representative and taken randomly.

Because the data sample is drawn from a population[6] of values, we make inferences about the population from certain sample statistics and from assumptions about the way the population of values is distributed. As shown in Figure 1, a sample histogram can aid in our attempt to picture the population distribution.

The normal distribution is ordinarily assumed to be the underlying distribution for repeatability or random error. When samples are taken, the sample mean and the sample standard deviation are computed and assumed to represent the mean and standard deviation of the population distribution. However, this equivalence is only approximate. To account for this, the Student's *t* distribution is used in place of the normal distribution to compute confidence limits around the sample mean.

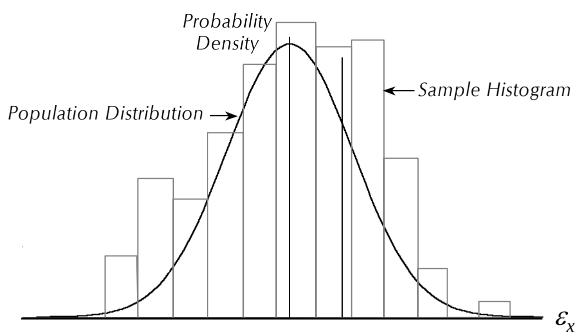


Figure 1. Repeatability Distribution

The sample mean, $\bar{x}$, is obtained by taking the average of the sampled values. The average value is computed by summing the values sampled and dividing them by the sample size, *n*.

$$\bar{x} = \frac{1}{n}\left(x_1 + x_2 + ... + x_n\right) = \frac{1}{n}\sum_{i=1}^{n} x_i \tag{6}$$

The sample mean can be thought of as an estimate of the value that we expect to get when we make a measurement. This "expectation value" is called the population mean, which is expressed by the symbol μ.

The sample standard deviation provides an estimate of the population standard deviation. The sample standard deviation, s_x, is computed by taking the square root of the sum of the squares of sampled deviations from the mean divided by the sample size minus one.

$$s_x = \sqrt{\frac{1}{n-1}\sum_{i=1}^{n}\left(x_i - \bar{x}\right)^2} \tag{7}$$

---

[5] Data may also be comprised of sampled mean values or sampled cells. The computed statistics vary slightly depending on the sample type.

[6] In statistics, a population is the total set of possible values for a random variable under consideration.

The value n-1 is the degrees of freedom for the estimate, which signifies the number of independent pieces of information that go into computing the estimate. Absent any systematic influences during sample collection, the sample standard deviation will approach its population counterpart as the sample size or degrees of freedom increases. The degrees of freedom for an uncertainty estimate are useful for establishing confidence limits and other decision variables.

The sample standard deviation provides an estimate of the repeatability or random error population standard deviation, $\sigma_{\varepsilon_{x,ran}}$. As discussed previously, the standard deviation of an error distribution is equal to the square root of the distribution variance

$$\sigma_{\varepsilon_{x,ran}} = \sqrt{\mathrm{var}(\varepsilon_{x,ran})} \tag{8}$$

Therefore, the sample standard deviation provides an estimate of the uncertainty due to repeatability or random error.[7]

$$u_{\varepsilon_{x,ran}} \cong s_x \tag{9}$$

If the objective of the uncertainty analysis is to characterize a given single measurement performed under specific circumstances, as in developing a statement of capability, then equation (9) should be used.

If the standard deviation in the mean value of the estimate is intended to represent the uncertainty in the mean value due to repeatability or random error, then the uncertainty in the mean value can be estimated to be equal to the standard deviation divided by the square root of the sample size.

$$u_{\varepsilon_{\bar{x},ran}} \cong \frac{s_x}{\sqrt{n}} \tag{10}$$

Type B Estimates

With the exception of repeatability or random error, the uncertainty in each error source must be estimated heuristically from the containment limits, $\pm L$, containment probability, p, and the inverse probability density function for the error distribution , $\mathrm{F}^{-1}(p)$, as shown in equation (11).

$$u = \frac{L}{\mathrm{F}^{-1}(p)} \tag{11}$$

For example, if the measurement process error is normally distributed, then the uncertainty is computed from

$$u = \frac{L}{\Phi^{-1}\left(\frac{1+p}{2}\right)} \tag{12}$$

where $\Phi^{-1}()$ is the inverse normal distribution function.[8]

Containment limits may be taken from manufacturer tolerance limits, stated expanded uncertainties obtained from calibration records or certificates, or statistical process control limits. Containment probability can also be obtained from service history data, for example, as the number of observed in-tolerances, $n_{\text{in-tol}}$, divided by the number of calibrations, N.

[7] The uncertainty due to repeatability or random error in measurement is estimated from a sample of measurements taken over a time period short enough to eliminate variations due to systematic drift or other factors.

[8] The inverse normal distribution function can be found in statistics texts and in most spreadsheet programs.

$$C\% = 100\% \, \frac{n_{\text{in-tol}}}{N}$$

In equation (12), the degrees of freedom are assumed to be infinite. However, we *know* that heuristic estimates are not based on an "infinite" amount of knowledge. As with Type A uncertainty estimates, the degrees of freedom quantify the amount of information that goes into the Type B uncertainty estimate.

Therefore, if there is an uncertainty in the containment limits (e.g., $\pm L \pm \Delta L$) or the containment probability (e.g., $\pm p \pm \Delta p$), then it becomes imperative to estimate the degrees of freedom. Equation G.3 of the GUM provides a relationship for computing the degrees of freedom for a Type B uncertainty estimate

$$\nu \approx \frac{1}{2}\frac{u^2(x)}{\sigma^2\left[u(x)\right]} \approx \frac{1}{2}\left[\frac{\Delta u(x)}{u(x)}\right]^{-2} \tag{13}$$

where $\sigma^2[u(x)]$ is the variance in the uncertainty estimate, $u(x)$, and $\Delta u(x)$ is the uncertainty in the uncertainty estimate.[9] Hence, the degrees of freedom for a Type B estimate are inversely proportional to the square of the ratio of the uncertainty in the uncertainty divided by the uncertainty.

While this approach is intuitively appealing, the GUM offers no advice about how to determine $\sigma^2[u(x)]$ or $\Delta u(x)$. Since the publication of the GUM, a methodology for determining $\sigma^2[u(x)]$ and computing the degrees of freedom for Type B estimates has been developed.[10]

Although the normal distribution is most often used to estimate uncertainty, other distributions also have limited applicability. As previously discussed, many of these distributions are described by minimum bounding limits, $\pm a$ and 100% containment probability (i.e., $p = 1$). Uncertainty equations for selected distributions are summarized in Table 1.

GUM METHOD

The GUM method evaluates and combines the variances of each measurement process error distribution. The combined uncertainty is computed by taking the square-root of the combined variance. The effective degrees of freedom for the combined uncertainty estimate are computed using the Welch-Satterthwaite relation. A Taylor Series approximation is employed for analyzing multivariate measurements.

Combined Uncertainty – Direct Measurements

The variance addition is employed to obtain a method for combining uncertainties that accounts for possible correlation between errors. To illustrate variance addition, let us consider the error in measurement of a quantity x that involves errors due to repeatability and measuring equipment bias, ε_{rep} and ε_{bias}.

$$\varepsilon_x = \varepsilon_{bias} + \varepsilon_{rep}$$

From variance addition, the uncertainty in ε_x is obtained from

$$\begin{aligned} u_{\varepsilon_x} &= \sqrt{\text{var}(\varepsilon_x)} = \sqrt{\text{var}(\varepsilon_{bias} + \varepsilon_{rep})} \\ &= \sqrt{\text{var}(\varepsilon_{bias}) + \text{var}(\varepsilon_{rep}) + 2\,\text{cov}(\varepsilon_{bias}, \varepsilon_{rep})} \\ &= \sqrt{u_{\varepsilon_{bias}}^2 + u_{\varepsilon_{rep}}^2 + 2\,\text{cov}(\varepsilon_{bias}, \varepsilon_{rep})} \end{aligned} \tag{14}$$

---

[9] This equation assumes that the underlying error distribution is normal.

[10] Castrup, H.: "Estimating Category B Degrees of Freedom," presented at the 2000 Measurement Science Conference, January 21, 2001.

where $u_{\varepsilon_{bias}}$ and $u_{\varepsilon_{rep}}$ are the measurement process uncertainties and $\text{cov}(\varepsilon_{bias}, \varepsilon_{rep})$ is the covariance between ε_{bias} and ε_{rep}.

The covariance of two random variables is a statistical assessment of their mutual dependence. Covariances are rarely used explicitly. Instead, we use the correlation coefficient, $\rho_{\varepsilon_{bias}\varepsilon_{rep}}$, which is defined as

$$\rho_{\varepsilon_{bias}\varepsilon_{rep}} = \frac{\text{cov}(\varepsilon_{bias}, \varepsilon_{rep})}{\sqrt{\text{var}(\varepsilon_{bias})\,\text{var}(\varepsilon_{rep})}} = \frac{\text{cov}(\varepsilon_{bias}, \varepsilon_{rep})}{u_{\varepsilon_{bias}} u_{\varepsilon_{rep}}} \tag{15}$$

Equation (14) can then be expressed as

$$u_{\varepsilon_x} = \sqrt{u_{\varepsilon_{bias}}^2 + u_{\varepsilon_{rep}}^2 + 2\rho_{\varepsilon_{bias},\varepsilon_{rep}} u_{\varepsilon_{bias}} u_{\varepsilon_{rep}}} \; . \tag{16}$$

The correlation coefficient is a dimensionless number ranging in value from -1 to 1. If the two errors are statistically independent, then $\rho_{\varepsilon_{bias}\varepsilon_{rep}} = 0$ and

$$u_{\varepsilon_x} = \sqrt{u_{\varepsilon_{bias}}^2 + u_{\varepsilon_{rep}}^2} \; .$$

If the two error sources are strongly correlated, then $\rho_{\varepsilon_{bias}\varepsilon_{rep}} = 1$ and

$$u_{\varepsilon_x} = \sqrt{u_{\varepsilon_{bias}}^2 + u_{\varepsilon_{rep}}^2 + 2u_{\varepsilon_{bias}} u_{\varepsilon_{rep}}} = u_{\varepsilon_{bias}} + u_{\varepsilon_{rep}} \; .$$

When two errors are strongly correlated and compensate for one another, then $\rho_{\varepsilon_{bias}\varepsilon_{rep}} = -1$ and

$$u_{\varepsilon_x} = \sqrt{u_{\varepsilon_{bias}}^2 + u_{\varepsilon_{rep}}^2 - 2u_{\varepsilon_{bias}} u_{\varepsilon_{rep}}} = \left| u_{\varepsilon_{bias}} - u_{\varepsilon_{rep}} \right| \; .$$

There typically aren't any correlations between measurement process errors for directly measured quantity.[11] Therefore, it is safe to assume that $\rho_{\varepsilon_{bias}\varepsilon_{rep}} = 0$.

Combined Uncertainty – Multivariate Measurements

For multivariate measurements, a more general equation is used for the combined uncertainty

$$u_c = \sqrt{\sum_{r=1}^{k} a_r^2 u_{\varepsilon_r}^2 + 2\sum_{r=1}^{k-1}\sum_{q=r+1}^{k} a_r a_q \sum_{i=1}^{l}\sum_{j=1}^{m} \rho_{\varepsilon_{r,i}\varepsilon_{q,j}} u_{\varepsilon_{r,i}} u_{\varepsilon_{q,j}}} \; . \tag{17}$$

One can surmise from equation (17) that uncertainties are not always combined using the root sum square (RSS) method. In the first term in equation (17), u_{ε_r} represents the combined uncertainty for each directly measured quantity and a_r accounts for corresponding sensitivity coefficients. The second term accounts for possible cross-correlations between measurement process uncertainties for the *r*th and *q*th measured quantities, $u_{\varepsilon_{r,i}}$ and $u_{\varepsilon_{q,j}}$.

[11] However, cross-correlations between measurement process errors may exist for multivariate measurements.

Effective Degrees of Freedom

Equation G.2b of the GUM provides the Welch-Satterthwaite relation as a means of computing the effective degrees of freedom for a combined uncertainty estimate, u_c.

$$\nu_{u_c} = \frac{u_c^4}{\sum_{i=1}^{n} \frac{u_i^4}{\nu_i}} \tag{18}$$

In the Welch-Satterthwaite relation, u_c is computed assuming that no error correlations exist and that the sensitivity coefficients for measurement process uncertainties, u_i, are all equal to one. Another underlying assumption of Welch-Satterthwaite relation is that the constituent errors are normally distributed.[12]

Confidence Limits

The combined uncertainty, u_c, and effective degrees of freedom, ν_{u_c}, can be used to establish confidence limits that contain the true value, μ (estimated by the mean value $\bar{x}$), with some specified confidence level or probability, p. Confidence limits are expressed as

$$\bar{x} - t_{\alpha/2,\nu} u_c \le \mu \le \bar{x} + t_{\alpha/2,\nu} u_c \tag{19}$$

where the multiplier is the t-statistic, $t_{\alpha/2,\nu}$, and $\alpha = 1 - p$. Values for $t_{\alpha/2,\nu}$, listed in Table 2, are obtained from the percentiles of the probability density function for the Student's t distribution.

As seen from equation (19), the width of the confidence limits or interval is dependent on three factors:

1. the combined uncertainty
2. the confidence level
3. the degrees of freedom.

Table 2. Values of the t-statistic for the Student's t Distribution[13]

Degrees of Freedom	$\alpha = 1 - p$ where p is the probability (fraction)							
ν	0.400	0.200	0.100	0.050	0.025	0.010	0.005	0.0005
1	0.325	1.3764	3.078	6.314	12.706	31.821	63.657	636.619
2	0.289	1.0607	1.886	2.920	4.303	6.965	9.925	31.598
3	0.277	0.9785	1.638	2.353	3.182	4.541	5.841	12.924
4	0.271	0.9410	1.533	2.132	2.776	3.747	4.604	8.610
5	0.267	0.9195	1.476	2.015	2.571	3.365	4.032	6.869
6	0.265	0.9057	1.440	1.943	2.447	3.145	3.707	5.959
7	0.263	0.8960	1.415	1.895	2.365	2.998	3.499	5.408
8	0.262	0.8889	1.397	1.860	2.306	2.896	3.355	5.041
9	0.261	0.8834	1.383	1.833	2.262	2.821	3.250	4.781
10	0.260	0.8791	1.372	1.812	2.228	2.764	3.169	4.587
11	0.260	0.8755	1.363	1.796	2.201	2.718	3.106	4.437
12	0.259	0.8726	1.356	1.782	2.179	2.681	3.055	4.318
13	0.259	0.8702	1.350	1.771	2.160	2.650	3.012	4.221
14	0.258	0.8681	1.345	1.761	2.145	2.624	2.977	4.140
15	0.258	0.8662	1.341	1.753	2.131	2.602	2.947	4.073

[12] More generalized forms of the Welch-Satterthwaite relation are derived in reference [5] for uncorrelated and correlated errors, as well as for multivariate measurements where the sensitivity coefficients may not equal one.

[13] From *CRC Standard Mathematical Tables*, 28th Edition, CRC Press, Inc., 2000.

Degrees of Freedom	α = 1- p where p is the probability (fraction)							
ν	0.400	0.200	0.100	0.050	0.025	0.010	0.005	0.0005
16	0.258	0.8647	1.337	1.746	2.120	2.583	2.921	4.015
17	0.257	0.8633	1.333	1.740	2.110	2.567	2.898	3.965
18	0.27	0.8620	1.330	1.734	2.101	2.552	2.878	3.922
19	0.257	0.8610	1.328	1.729	2.093	2.539	2.861	3.883
20	0.257	0.8600	1.325	1.725	2.086	2.528	2.845	3.850
21	0.257	0.8591	1.323	1.721	2.080	2.518	2.831	3.819
22	0.256	0.8583	1.321	1.717	2.074	2.508	2.819	3.792
23	0.256	0.8575	1.319	1.714	2.069	2.500	2.807	3.767
24	0.256	0.8569	1.318	1.711	2.064	2.492	2.797	3.745
25	0.256	0.8562	1.316	1.708	2.060	2.485	2.787	3.725
26	0.256	0.8557	1.315	1.706	2.056	2.479	2.779	3.707
27	0.256	0.8551	1.314	1.703	2.052	2.473	2.771	3.690
28	0.256	0.8546	1.313	1.701	2.048	2.467	2.763	3.674
29	0.256	0.8542	1.311	1.699	2.045	2.462	2.756	3.659
30	0.256	0.8538	1.310	1.697	2.042	2.457	2.750	3.646
40	0.256	0.8507	1.303	1.684	2.021	2.423	2.704	3.551
60	0.254	0.8477	1.296	1.671	2.000	2.390	2.660	3.460
120	0.254	0.8446	1.289	1.658	1.980	2.358	2.617	3.373
∞	0.253	0.8416	1.282	1.645	1.960	2.326	2.576	2.291

Note: For two-sided confidence limits, the column in Table 2 with the value of $\alpha/2$ must be selected. For example, given a probability of 0.95 and degrees of freedom of 20, $\alpha/2 = 0.025$ and $t_{0.025,20} = 2.086$.

The major assumption is that the combined error follows a normal (infinite degrees of freedom) or Student's t distribution[14] (finite degrees of freedom) results from the *central limit theorem*, which demonstrates that the combined error distribution converges toward the normal distribution as the number of constituent errors increases, regardless of their underlying distributions. This is illustrated in Figure 1, where the combined probability distribution for three errors should take on a normal or Gaussian shape, regardless of the shape of the individual error distributions. The actual shape of combined error distributions will be discussed further for the convolution and Monte Carlo methods.

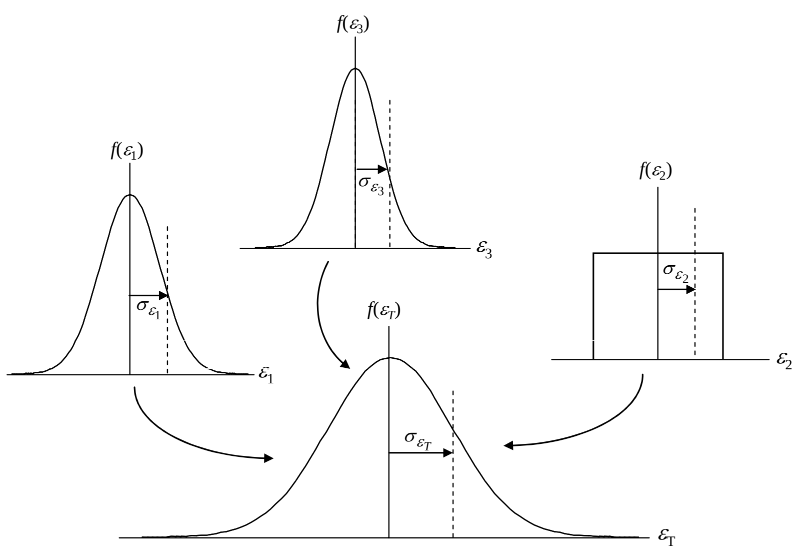


Figure 1. Combined Measurement Error Distribution

[14] The Student's t distribution is a symmetric distribution that approaches the normal distribution as the degrees of freedom approach infinity.

Expanded Uncertainty

Expanded uncertainty is defined in the GUM as "the 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." In other words, the expanded uncertainty is basically defined as a set of limits (± U) that are expected to contain the true value of the measurand.[15]

$$\bar{x} - U \le x_{true} \le \bar{x} + U \tag{20}$$

The expanded uncertainty is defined as $U = ku_c$ and is offered as an approximate confidence limit, in which a coverage factor, k, is used in place of the t-statistic.

$$\bar{x} - ku_c \le \mu \le \bar{x} + ku_c \tag{21}$$

In the application of equation (21), it is assumed that u_c has infinite degrees of freedom and the combined error is normally distributed. In such cases, the coverage factor is obtained from the last row of Table 2 for a given confidence level. It is also a common practice to round the value of k to the nearest whole number.

CONVOLUTION METHOD

If two or more errors are statistically independent, then the distribution for the combined error can be obtained by convolution. To illustrate the convolution method, let ε_x and ε_y be two statistically independent, continuously distributed measurement process errors with probability density functions $f(\varepsilon_x)$ and $g(\varepsilon_y)$, respectively. The distribution for the combined error

$$\varepsilon = \varepsilon_x + \varepsilon_y \tag{22}$$

can be obtained from the relation

$$h(\varepsilon) = \int_{-\infty}^{\infty} f(\varepsilon_x) g(\varepsilon - \varepsilon_x) d\varepsilon_x \ . \tag{23}$$

The convolution method is applicable for direct measurements where the measurement process errors are statistically independent (i.e., no error correlations). Several examples showing how two or more error distributions are convolved are given in reference [4]. A few examples are provided here for illustration.

Convolved Uniform Distributions

First, we consider two statistically independent, uniformly distributed errors ε_1 and ε_2 with the following probability density functions

$$f(\varepsilon_1) = \begin{cases} \dfrac{1}{2a}, & -a \le \varepsilon_1 \le a \\ 0, & \text{otherwise}, \end{cases} \tag{24}$$

and

$$g(\varepsilon_2) = \begin{cases} \dfrac{1}{2b}, & -b \le \varepsilon_2 \le b \\ 0, & \text{otherwise}. \end{cases} \tag{25}$$

where ± a and ± b are the minimum bounding limits for the respective error distributions. Using equation (23), the convolved distribution is described by

[15] The term expanded uncertainty and uncertainty are often used interchangeably. This, of course, should be avoided because it can lead to incorrect inferences and miscommunications.

$$h(\varepsilon) = \begin{cases} \frac{1}{4ab}(a+b+\varepsilon)\,, & -(a+b) \le \varepsilon \le -(b-a) \\ \frac{1}{2b}\,, & -(b-a) \le \varepsilon \le b-a \\ \frac{1}{4ab}(a+b-\varepsilon)\,, & b-a \le \varepsilon \le b+a \\ 0\,. & \text{otherwise} \end{cases} \tag{26}$$

When the error distributions have different 100% containment limits ($b > a$), the combined error distribution takes on a trapezoidal shape, as shown in Figure 2.

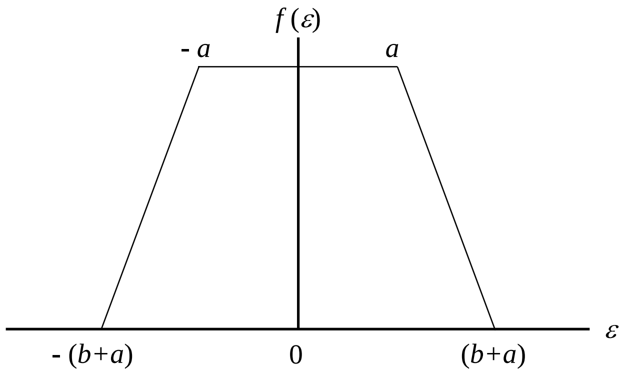


Figure 2. Convolved Distribution for Two Uniformly Distributed Errors, *b > a*.

When the error distributions have equal 100% containment limits ($b = a$), the combined error distribution becomes the familiar triangular distribution shown in Figure 3.

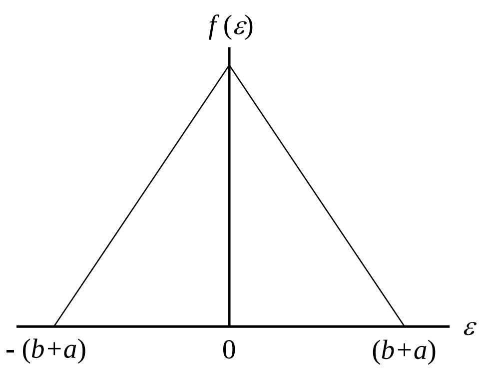


Figure 3. Convolved Distribution for Two Uniformly Distributed Errors, *b = a*.

Now, let us consider three statistically independent uniformly distributed errors ε_1, ε_2 and ε_3 with the following probability density functions

$$f(\varepsilon_1) = \begin{cases} \frac{1}{2a}, & -a \le \varepsilon_1 \le a \\ 0, & \text{otherwise}, \end{cases} \quad , \quad g(\varepsilon_2) = \begin{cases} \frac{1}{2b}, & -b \le \varepsilon_2 \le b \\ 0, & \text{otherwise}, \end{cases} \quad \text{and} \quad w(\varepsilon_3) = \begin{cases} \frac{1}{2c}, & -c \le \varepsilon_3 \le c \\ 0, & \text{otherwise}. \end{cases}$$

A number of combinations of relative sizes of the 100% minimum bounding limits $\pm a$, $\pm b$ and $\pm c$ are possible. One case, where $b > a$ and $c > a + b$, will be explicitly constructed. The probability density function for the convolved error distribution is broken up into six pieces:

$$h(\varepsilon) = \frac{1}{8abc}(c+b+a+\varepsilon)^2\ , \quad -(a+b+c) \le \varepsilon \le -(b-c)-a$$

$$h(\varepsilon) = \frac{1}{4ab}(a+b+\varepsilon)\ , \quad -(b-c)-a \le \varepsilon \le (b-c)-a$$

$$h(\varepsilon) = \frac{1}{8abc}\left[\frac{1}{2}(a+b+c+\varepsilon)^2 - (a+\varepsilon)^2 - (b-c)^2\right], \quad (b-c)-a \le (b+c)-a$$

$$h(\varepsilon) = \frac{1}{2a}\ , \quad (b+c)-a \le \varepsilon \le a-(b+c)$$

$$h(\varepsilon) = \frac{1}{8abc}\left[4bc - \frac{1}{2}(b+c-a+\varepsilon)^2\right], \quad a-(b+c) \le \varepsilon \le a-(b-c)$$

$$h(\varepsilon) = \frac{1}{8abc}(a+b+c-\varepsilon)^2\ , \quad (b-c)+a \le \varepsilon \le a+b+c\ .$$

The combined error distribution, shown in Figure 4, begins to look more like the normal distribution. Actually, its shape is that of the utility distribution.[16] The convolved distribution begins to look even more normally distributed in cases where $c > b$ and $b = a$ and for cases where $c = b = a$.

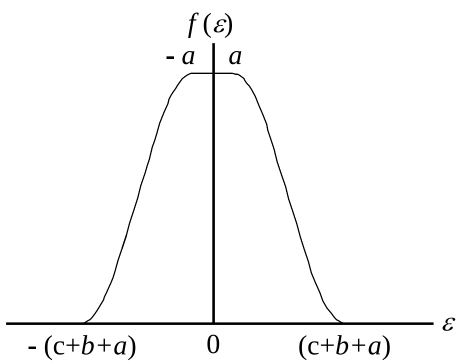


Figure 4. Convolved Distribution for Three Uniformly Distributed Errors, b > a and c > a + b

As previously stated, the uniform distribution is applicable for only a few measurement processes errors. For direct measurement scenarios, it is unlikely to have more than a couple of measurement process errors that follow a uniform distribution.[17] The primary purpose of convolving three uniform distributions is to illustrate that the central limit theorem holds true, regardless of the shape of the individual error distributions.

Convolved Uniform and Normal Distributions

The normal distribution applies to a wide variety of measurement process errors, so let us consider two statistically independent errors, one that is uniformly distributed and the other normally distributed. In this example, the 100% minimum bounding limits for the uniform distribution are ± 3.5 and the standard deviation of the normal distribution is 1.

Consequently, the standard deviation of the uniform distribution is approximately twice that of the normal distribution ($\sigma_U \cong 2\sigma_N$). The convolved distribution is shown in Figure 5. If additional normally distributed measurement process errors were evaluated, the convolved error distribution would exhibit a more Gaussian or normal shape.

[16] See reference [4] for a description of the utility function and its probability density function.

[17] Two uniformly distributed errors may result from the digital resolution of a unit under test parameter and a measurement reference.

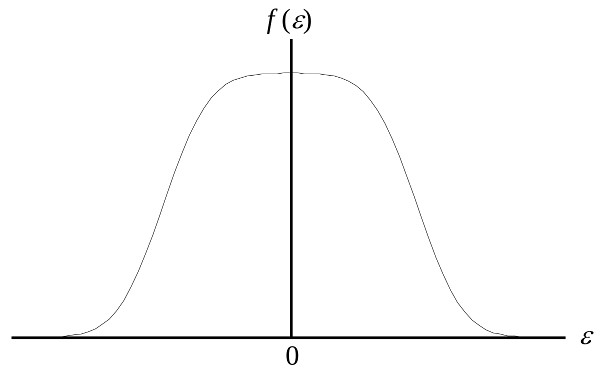


Figure 5. Convolved Distribution for Uniformly and Normally Distributed Errors, $\sigma_U \cong 2\sigma_N$

Confidence Limits

The confidence limits are computed via numerical integration of the combined error distribution for a given percent confidence level (% *C*) or probability ($p = C/100$). In this case, numerical iteration using a bisection method can be employed.[18]

Expanded Uncertainty

The expanded uncertainty, $\pm U$, is computed from the standard deviation of the convolved error distribution and the specified coverage factor (usually $k = 2$).

$$\pm U = \pm\left(k \times u_c\right) \tag{27}$$

MONTE CARLO METHOD

Monte Carlo simulation is another method used to combine error distributions. As with the previous two methods, the appropriate distribution must be selected for each measurement process error. The standard deviation or the error containment limits and containment probability must also be known for each error distribution. The Monte Carlo method is also applicable for multivariate measurement scenarios.

The Monte Carlo method employs repeated computation of random or pseudo-random numbers to simulate and combine deviates for individual error distributions. The combined error distribution is obtained by adding the simulated data points for all of the distributions. The Monte Carlo method can be used to analyze measurements with correlated errors. Correlation coefficients equal to unity can be easily accommodated, but generalized cases of statistically correlated errors are considerably more difficult to implement.

Simulated Uniform and Normal Distributions

We will evaluate the uniform and normal distributions previously described, in which the standard deviation of the uniform distribution is approximately twice that of the normal distribution. Because Monte Carlo simulation involves the generation of pseudo-random values, the results are subject to statistical fluctuations, as seen in each distribution plot.

The scatter in the data points is affected by the number of simulated values (or trials) that are generated[19] and the bin size used to sort them. The frequency or probability distribution is determined from the number of values that fall in each bin. Bin size is usually set equal to the maximum expected value minus the minimum expected value divided by the number of bins.

The larger the number of simulated values (or trials), the better the agreement between the data points and the underlying distribution. The simulated uniform and normal distributions shown in Figures 6 and 7, respectively, were each generated from 10,000 simulated trials using the MS Excel Random Number Generator analysis tool.

---

[18] *Numerical Recipes in Fortran*, 2nd Edition, Cambridge University Press, 1992.

[19] A minimum of 10,000 simulated deviates are required for each error distribution to obtain meaningful results.

The data were then sorted into 60 bins to obtain their relative frequency of occurrence.

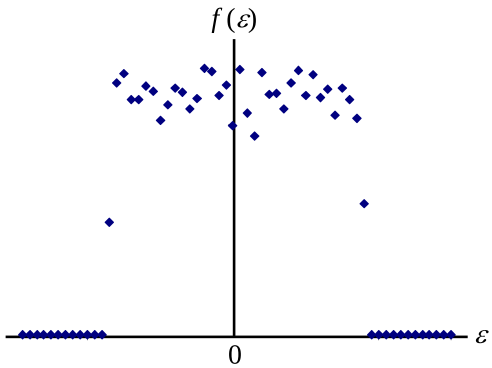


Figure 6. Simulated Uniform Distribution

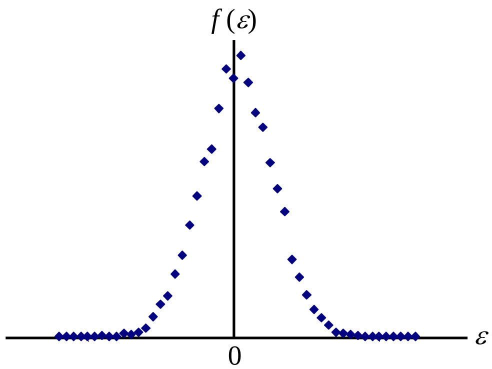


Figure 7. Simulated Normal Distribution

Combined Uniform and Normal Distribution

The combined distribution for the simulated uniform and normal shown in Figures 6 and 7 was obtained by adding pairs of simulated data points. To generate the necessary frequency data, the number of bins was increased from 60 to 80 to maintain the same bin size that was used for the uniform and normal distributions. The combined distribution shown in Figure 8 looks very similar to the convolved distribution shown in Figure 5. This is expected, since the same standard deviations were used for the underlying uniform and normal distributions (i.e.,$\sigma_U \cong 2\sigma_N$).

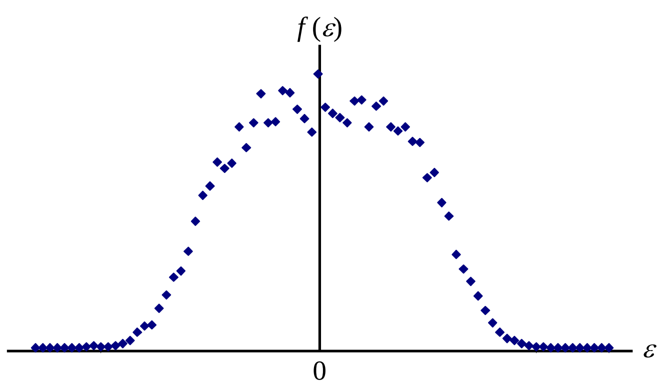


Figure 8. Combined Uniform and Normal Distributions, $\sigma_U \cong 2\sigma_N$

Consequently, the standard deviation for the combined distribution should be in close agreement with the standard deviation of the convolved distribution. In both analyses, the standard deviation of the normal distribution is $\sigma_N = 1$ and the standard deviation of the uniform distribution is

$$\sigma_U = \frac{3.5}{\sqrt{3}} = 2.0207 .$$

The standard deviation of the convolved distribution is computed to be 2.255 and the standard deviation for the combined distribution shown in Figure 8 is computed to be 2.252. For comparison, the uncertainty (i.e., standard deviation) for the combined error was also computed using the GUM method and its value was computed to be 2.255.

Confidence Limits

The confidence limits are computed via numerical integration of the combined error distribution for a given percent confidence level (% C) or probability ($p = C/100$). In this case, numerical iteration using a bisection method can be employed.

Expanded Uncertainty

The expanded uncertainty, $\pm U$, is computed from the standard deviation of the combined error distribution and the specified coverage factor (usually $k = 2$), as shown in equation (27).

DIRECT MEASUREMENT SCENARIOS

Four direct measurement scenarios described in reference [7] were analyzed using the GUM, convolution and Monte Carlo methods.[20] The GUM analyses for each scenario are summarized in the following sections. The GUM analysis results are then compared to the combined uncertainties and confidence limits obtained using the convolution and Monte Carlo methods.

Scenario 1: Calibration of Mass using Precision Balance

This measurement scenario consists of calibrating a 30 gm mass with a precision balance. In this scenario, the following measurement process errors are taken into account:

- Bias in the precision balance, $e_{MTE,b}$.
- Error due to the digital resolution of the balance, $e_{MTE,res}$.
- Environmental factors error resulting from the buoyancy correction, e_{env}.

The combined measurement error is

$$\varepsilon_{cal} = e_{MTE,b} + e_{MTE,res} + e_{env}$$

where

$$e_{env} = e_{UUT,env} = c_1 e_{\rho_{air}} + c_2 e_{\rho_{UUT}}$$

and $e_{\rho_{air}}$ and $e_{\rho_{UUT}}$ are the errors in the air and UUT densities, respectively.

In this analysis, the error in the air density and the error in the density of the UUT mass are considered to be uncorrelated. The coefficients c_1 and c_2 are sensitivity coefficients that determine the relative contribution of $e_{\rho_{air}}$ and $e_{\rho_{UUT}}$ to e_{env}. The sensitivity coefficients are defined and computed in reference [7]. The probability distributions, limits, confidence levels and uncertainties for each measurement process error are summarized in Table 3.

[20] For simplicity, repeatability is not included as a measurement process error in these analyses.

Table 3. Summary of Scenario 1 Uncertainty Estimates

Error Source	Error Limits	Confidence Level (%)	Error Distribution	Degrees of Freedom	Analysis Type	Standard Uncertainty	Sensitivity Coefficient
$e_{MTE,b}$	± 0.12 gm	95.00	Normal	Infinite	B	6.12×10^{-2} gm	1
e_{res}	± 0.005 gm	100.00	Uniform	Infinite	B	2.9×10^{-3} gm	1
$e_{\rho_{air}}$	± 3.6×10^{-5} gm/cm^3	95.00	Normal	Infinite	B	1.84×10^{-6} gm/cm^3	-0.18 cm^3
$e_{\rho_{UUT}}$	± 0.15 gm/cm^3	95.00	Normal	Infinite	B	0.077 gm/cm^3	-5.1×10^{-4} cm^3

Using the GUM method, the combined uncertainty equation is

$$u_{cal} = \sqrt{u_{MTE,b}^2 + u_{MTE,res}^2 + c_1^2 u_{\rho_{air}}^2 + c_2^2 u_{\rho_{UUT}}^2} = \sqrt{u_{MTE,b}^2 + u_{MTE,res}^2 + \left(c_1 u_{\rho_{air}}\right)^2 + \left(c_2 u_{\rho_{UUT}}\right)^2}\ .$$

Using the data in Table 3, the combined uncertainty is computed to be

$$\begin{aligned} u_{cal} &= \sqrt{\left(6.12\times10^{-2}\right)^2 + \left(2.9\times10^{-3}\right)^2 + \left(-0.18\times1.84\times10^{-6}\right)^2 + \left(-5.1\times10^{-4}\times0.077\right)^2}\ \text{gm} \\ &= \sqrt{3.75\times10^{-3} + 8.41\times10^{-6} + 1.81\times10^{-10} + 1.08\times10^{-13}}\ \text{gm} \\ &= \sqrt{3.75\times10^{-3}}\ \text{gm} = 6.12\times10^{-2}\ \text{gm}. \end{aligned}$$

Using the Welch-Satterthwaite relation, the combined degrees of freedom are computed to be infinite.

Scenario 2: Calibration of an Analog Micrometer using a Gage Block

This measurement scenario consists of calibrating an analog micrometer with a 10 mm gage block reference. In the micrometer calibration scenario, we must account for the following measurement process errors:

- Bias in the value of the 10 mm gage block length, $e_{MTE,b}$.
- Error associated with the analog resolution of the micrometer, $e_{UUT,res}$.
- Bias resulting from the operator's use of the micrometer to measure the gage block, $e_{UUT,op}$.
- Environmental factors error resulting from the thermal expansion correction, e_{env}.

The combined measurement error is

$$\varepsilon_{cal} = e_{MTE,b} + e_{UUT,res} + e_{UUT,op} + e_{env}$$

where

$$e_{env} = c_{\Delta T}\, e_{\Delta T}$$

and $c_{\Delta T}$ is the sensitivity coefficient and $e_{\Delta T}$ is the error due to the environmental temperature variation. The sensitivity coefficient, $c_{\Delta T}$, is defined and computed in reference [7].

Using the GUM method, the combined uncertainty equation is

$$u_{cal} = \sqrt{u_{MTE,b}^2 + u_{UUT,res}^2 + u_{UUT,op}^2 + c_{\Delta T}^2 u_{\Delta T}^2}\ .$$

The probability distributions, limits, confidence levels and standard uncertainties for each measurement process error are summarized in Table 4.

Table 4. Summary of Scenario 2 Uncertainty Estimates

Error Source	Error Limits	Conf. Level (%)	Error Distribution	Degrees of Freedom	Analysis Type	Standard Uncertainty	Sensitivity Coefficient
$e_{MTE,b}$	+ 0.18, -0.13 μm	90.00	Lognormal	Infinite	B	0.09 μm	1
$e_{UUT,res}$	± 5.0 μm	95.00	Normal	Infinite	B	2.6 μm	1
$e_{UUT,op}$	± 5.0 μm	95.00	Normal	Infinite	B	2.6 μm	1
$e_{\Delta T}$	± 1 °C	95.00	Normal	Infinite	B	0.51 °C	-5.9×10^{-2} μm/°C

Using the data in Table 4, the combined uncertainty is computed to be

$$\begin{aligned} u_{cal} &= \sqrt{(0.09)^2 + (2.6)^2 + (2.6)^2 + \left(-5.9\times10^{-2}\times0.51\right)^2}\ \mu\text{m} \\ &= \sqrt{0.0081 + 6.76 + 6.76 + 0.0009}\ \mu\text{m} \\ &= \sqrt{13.53}\ \mu\text{m} = 3.68\ \mu\text{m} \end{aligned}$$

Using the Welch-Satterthwaite relation, the combined degrees of freedom are computed to be infinite.

Scenario 3: Calibration of an End Gauge using a Comparator

This measurement scenario consists of calibrating an end gauge, with a nominal length of 50 mm, using an end gauge standard of the same nominal length. The calibration process consists of measuring and recording the difference between the two end gauges using a comparator apparatus.

In this calibration scenario, the following measurement process errors are taken into account:

- Bias in the value of the 50 mm end gage standard length, $e_{MTE,b}$.
- Bias of the comparator, $e_{c,b}$
- Digital Resolution error for the comparator, e_{res}.
- Operator bias, e_{op}.
- Environmental factors error resulting from the thermal expansion correction, e_{env}.

The combined measurement error is

$$\varepsilon_{cal} = (\varepsilon_{UUT,m} - \varepsilon_{MTE,m}) + e_{op} - e_{MTE,b}$$

where

$$\varepsilon_{MTE,m} = e_{c,b} + e_{MTE,res} + e_{MTE,env}$$

and

$$\varepsilon_{UUT,m} = e_{c,b} + e_{UUT,res} + e_{UUT,env}\,.$$

The expression for ε_{cal} can be rewritten as

$$\varepsilon_{cal} = e_{res} + e_{env} + e_{op} - e_{MTE,b}$$

where

$$e_{res} = e_{UUT,res} - e_{MTE,res}$$

$$e_{env} = e_{UUT,env} - e_{MTE,env} = c_1 e_{\Delta T} + c_2 e_{\alpha}$$

and $e_{\Delta T}$ is the error in the correction temperature and e_{α} is the error in the coefficient of thermal expansion for the UUT and reference ring gauges.

In this analysis, the correction temperature error and the coefficient of thermal expansion error are considered to be uncorrelated. The coefficients c_1 and c_2 are sensitivity coefficients that determine the relative contribution of $e_{\Delta T}$ and e_{α} to the environmental correction error e_{env}. The sensitivity coefficients are defined and computed in reference [7].

The resolution uncertainty for the UUT and MTE are equal to the resolution uncertainty of the comparator, $u_{UUT,res}$ = $u_{MTE,res}$ = $u_{c,res}$. In addition, the resolution error for the UUT and MTE are uncorrelated, so that ρ_{res} = 0. Using the GUM method, the combined uncertainty equation is

$$\begin{aligned} u_{cal} &= \sqrt{2u_{c,res}^2 + c_1^2 u_{\Delta T}^2 + c_2^2 u_{\alpha}^2 + 2c_1 c_2 u_{\Delta T} u_{\alpha} + u_{op}^2 + u_{MTE,b}^2} \\ &= \sqrt{2u_{c,res}^2 + \left(c_1 u_{\Delta T} + c_2 u_{\alpha}\right)^2 + u_{op}^2 + u_{MTE,b}^2} \end{aligned}.$$

The probability distributions, limits, confidence levels and uncertainties for each measurement process error are summarized in Table 5.

Table 5. Summary of Scenario 3 Uncertainty Estimates

Error Source	Error Limits	Conf. Level (%)	Error Distribution	Degrees of Freedom	Analysis Type	Standard Uncertainty (nm)	Sensitivity Coefficient
$e_{c,res}$	± 1 nm	100.0	Uniform	Infinite	B	0.577 nm	1
$e_{\Delta T}$	± 0.5 °C	95.00	Normal	Infinite	B	0.255 °C	2.47×10^{-3} nm/°C
e_{α}	± 0.5×10^{-6} /°C	95.00	Normal	Infinite	B	0.255×10^{-6} /°C	- 21.5 °C•nm
e_{op}	± 5 nm	95.00	Normal	Infinite	B	2.55 nm	1
$e_{MTE,b}$		95.00	Normal	Infinite	A,B	25 nm	1

Using the data in Table 5, the combined uncertainty is computed to be

$$\begin{aligned} u_{cal} &= \sqrt{2\times\left(0.577\right)^2 + (2.47\times10^{-3}\times0.255)^2 + (21.5\times0.255\times10^{-6})^2 + \left(2.55\right)^2 + \left(25\right)^2} \text{ nm} \\ &= \sqrt{0.67 + 3.97\times10^{-7} + 3.01\times10^{-11} + 6.50 + 625} \text{ nm} \\ &= \sqrt{632.2} \text{ nm} = 25.1 \text{ nm} \end{aligned}$$

Using the Welch-Satterthwaite relation, the combined degrees of freedom are computed to be infinite.

Scenario 4: Calibration of a Digital Thermometer

This measurement scenario consists of calibrating a digital thermometer at 100 °C using an oven and an analog temperature reference. The oven temperature is adjusted using its internal temperature probe and the readings from the thermometer and temperature reference are recorded.

In the thermometer calibration scenario, the following measurement process errors must be taken into account:

- Bias of the temperature reference, $e_{MTE,b}$.
- Analog resolution error for the temperature reference, $e_{MTE,res}$.
- Digital resolution error for the thermometer, $e_{UUT,res}$.
- Error due to the non-uniformity of the oven temperature, e_{env}.

In this analysis, the short-term effect of oven stability is not included as part of the environmental factors error, e_{env}.

The combined measurement error is

$$\varepsilon_{cal} = (e_{UUT,res} - e_{MTE,res}) - e_{MTE,b} + e_{env} .$$

Using the GUM method, the combined uncertainty equation is

$$u_{cal} = \sqrt{u_{MTE,b}^2 + u_{UUT,res}^2 + u_{MTE,res}^2 + u_{env}^2} .$$

The probability distributions, limits, confidence levels and standard uncertainties for each measurement process error are summarized in Table 6.

Table 6. Summary of Scenario 4 Uncertainty Estimates[21]

Error Source	Error Limits (°C)	Confidence Level (%)	Error Distribution	Degrees of Freedom	Analysis Type	Standard Uncertainty (°C)	Sensitivity Coefficient
$e_{MTE,b}$			Normal	infinite	A,B	0.02	1
$e_{UUT,res}$	± 1	100.00	Uniform	infinite	B	0.577	1
$e_{MTE,res}$	± 0.25	95.00	Normal	infinite	B	0.128	1
e_{env}	± 1	95.00	Normal	Infinite	B	0.51	1

Using the data in Table 6, the combined uncertainty is computed to be

$$\begin{aligned} u_{cal} &= \sqrt{(0.02)^2 + (0.577)^2 + (0.128)^2 + (0.51)^2} \text{ °C} \\ &= \sqrt{0.0004 + 0.3329 + 0.0163 + 0.2603} \text{ °C} \\ &= \sqrt{0.6099} \text{ °C} = 0.78 \text{ °C} \end{aligned}$$

Using the Welch-Satterthwaite relation, the combined degrees of freedom are computed to be infinite.

Comparison to Convolution and Monte Carlo Analysis Results

The four direct measurement scenarios were also analyzed using the convolution and Monte Carlo methods. The convolution analyses were conducted using the Convolver application developed by Dr. H. Castrup.[22] The Monte Carlo analyses were conducted using the MS Excel Random Number Generator analysis tool. Each error distribution was generated from 10,000 simulated trials that were sorted into 100 bins to obtain their relative frequency of occurrence.

The combined uncertainties and confidence limits computed using the GUM, convolution and Monte Carlo methods are summarized in Table 7. A couple of extra decimal digits were included to highlight any differences in the computed values. Of the three methods, convolution is considered to be the most rigorous approach for obtaining combined uncertainty estimates and confidence limits for the four direct measurement scenarios under consideration.

The combined uncertainties obtained from the GUM and Monte Carlo methods are in very close agreement with corresponding values obtained with the convolution method. The 95% confidence limits are also in close agreement for measurement scenarios 1 through 3. The 95% confidence limits computed with the GUM and Monte Carlo

[21] For the purposes of illustration, the error limits for $e_{UUT,res}$, $e_{MTE,res}$, and e_{env} were modified from those in reference [7] to increase the relative contribution of the uniformly distributed UUT resolution error.

[22] Convolver, ©2005-2009, Integrated Sciences Group, All Rights Reserved.

methods are comparable for measurement scenario 4, but are significantly higher than the limits obtained via convolution. The reason for this is evident from the combined error distributions shown in Figure 9, in which the GUM plot depicts a normal distribution with a standard deviation equal to the combined uncertainty. The Monte Carlo data follow the normal distribution more closely than the actual (convolved) error distribution. The smaller 95% confidence limits for the convolved error distribution are attributable to its distinctly different shape.

Table 7. Summary of Measurement Scenario Analysis Results

Meas. Scenario	GUM Combined Uncertainty	Convolution Combined Uncertainty	Monte Carlo Combined Uncertainty	GUM Confidence Limits (95%)	Convolution Confidence Limits (95%)	Monte Carlo Confidence Limits (95%)
1	0.0612 gm	0.0613 gm	0.0605 gm	± 0.120 gm	± 0.117 gm	± 0.115 gm
2	3.678 μm	3.678 μm	3.682 μm	± 7.213 μm	± 7.184 μm	± 7.140 μm
3	25.143 nm	25.136 nm	25.052 nm	± 49.280 nm	± 49.105 nm	± 49.000 nm
4	0.781 °C	0.781 °C	0.785 °C	± 1.531 °C	± 1.228 °C	± 1.470 °C

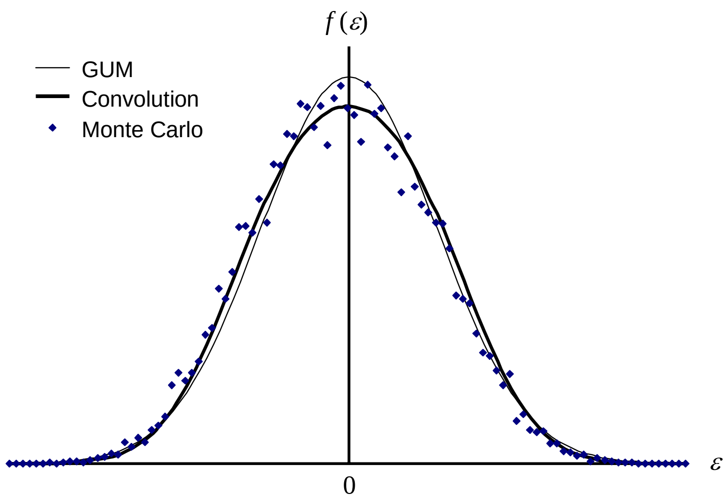


Figure 9. Combined Error Distributions for Measurement Scenario 4

CONCLUSIONS

The GUM, convolution and Monte Carlo methods were evaluated for estimating the uncertainty or standard deviation of a combined error distribution comprised of distributions for individual measurement processes errors. The three methods were used to compute and compare the combined uncertainty estimates and confidence limits for four direct measurement scenarios involving uncorrelated errors. The convolution method is considered to be the most rigorous approach for obtaining combined uncertainty estimates and confidence limits for these direct measurement scenarios.

In general, all three methods are considered to provide comparable analysis results. This is especially true in cases where the combined error distribution closely approaches that of the normal distribution. The combined uncertainties obtained from the GUM and Monte Carlo methods were found to be in very close agreement with the combined uncertainties obtained with the convolution method. Similarly, the 95% confidence limits computed from all three methods were in close agreement for measurement scenarios 1 through 3. For measurement scenario 4, GUM and Monte Carlo methods computed comparable 95% confidence limits, while the significantly different shape of the convolved error distribution resulted in smaller limits.

All three methods can be used to analyze direct and multivariate measurements in which all errors are uncorrelated. The GUM and Monte Carlo methods can also be used to analyze direct and multivariate measurements involving uncorrelated errors. The GUM and Monte Carlo methods can also accommodate correlated errors. However, applying the Monte Carlo method to cases were $\rho \neq 1$ is a challenging endeavor. Additional studies are needed to evaluate and compare these methods for multivariate measurements, with and without correlated errors.

REFERENCES

[1] ANSI/NCSL-Z540-2-1997 *U.S. Guide to the Expression of Uncertainty in Measurement*.

[2] NCSLI-RP-12 *Determining and Reporting Measurement Uncertainty*, Draft Re-write, February 2010.

[3] Hahn, G. J. and Shapiro, S. S.: Statistical Models in Engineering, John Wiley & Sons, Inc. 1967.

[4] Castrup, H.: "Selecting and Applying Error Distributions in Uncertainty Analysis," presented at the 2004 Measurement Science Conference in Anaheim, CA, January 16-17, 2004.

[5] Castrup, H.: “A Welch-Satterthwaite Relation for Correlated Errors,” presented at the 2010 Measurement Science Conference in Pasadena, CA, March 25-26, 2010.

[6] Press, W. H., et al.: *Numerical Recipes in Fortran*, 2nd Edition, Cambridge University Press, 1992.

[7] Castrup, H and Castrup, S.: "Uncertainty Analysis for Alternative Calibration Scenarios," presented at the 2008 NCSLI Conference in Orlando, FL, August 2008.