Error Distribution Variances and Other Statistics[1]

Dr. Howard Castrup
Integrated Sciences Group
1/15/2009

Introduction

This monograph was developed in response to customer inquiries concerning the expressions used to represent the standard deviations (standard uncertainties) of the various error probability distributions employed in ISG software and elsewhere. The following develops these standard deviations, showing in detail the steps involved. For each distribution, the steps may vary, but the basic procedure is to first develop an expression for the distribution variance σ^2, from which the standard deviation σ follows immediately.

Distribution Variance - The General Expression

For a random variable ε with probability distribution function (pdf) $f(\varepsilon)$ and mean μ, the variance is given by

$$\sigma^2 = \int_{-\infty}^{\infty} f(\varepsilon)(\varepsilon - \mu)^2 d\varepsilon \ . \tag{1}$$

For error distributions, μ is taken to be zero, and Eq. (1) reduces to

$$\sigma^2 = \int_{-\infty}^{\infty} f(\varepsilon)\varepsilon^2 d\varepsilon \ . \tag{2}$$

The Normal Distribution

The normal distribution is the "workhorse" of statistics and probability. It is usually assumed to be the underlying distribution for random variables. Indeed, the various tools we use in applying uncertainty estimates are nearly always based on the assumption that measurement errors are normally distributed, regardless of the distributions used to estimate the uncertainties themselves.

The pdf

The pdf for a normally distributed random variable ε with mean 0 and standard deviation σ is given by

$$f(\varepsilon) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\varepsilon^2/2\sigma^2} \ . \tag{3}$$

The variance σ^2 is the square of the standard deviation.

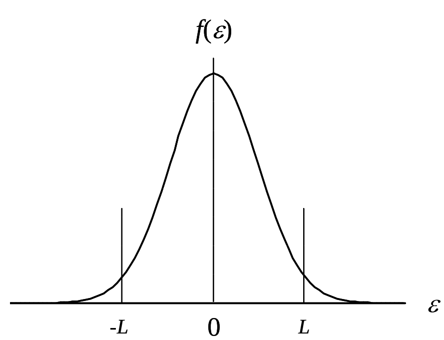

Figure 1. The Normal Distribution. Shown is a case where the error distribution mean value is zero. The limits ±*L* are approximate 95% confidence limits.

---

[1] Much of the material in this monograph is taken from Castrup, H., "Selecting and Applying Error Distributions in Uncertainty Analysis," presented at the Measurement Science Conference, Anaheim, 2004.

Applicability of the Normal Distribution

We usually assume a normal distribution. Why? The primary reason is because this is the distribution that either represents or approximates what we frequently see in the physical universe. It can be derived from the laws of physics for such phenomena as the diffusion of gases and is applicable to instrument parameters subject to random stresses of usage and handling. It is also often applicable to equipment parameters emerging from manufacturing processes.

An additional consideration applies to the distribution we should assume for a total error or deviation that is composed of constituent errors or deviations. There is a theorem called the *central limit theorem* that demonstrates that, even though the individual constituent errors or deviations may not be normally distributed, the combined error or deviation is approximately so.

An argument has been presented against the use of the normal distribution in cases where the variable of interest is restricted, i.e., where values of the variable are said to be bound by finite physical limits. This condition notwithstanding, the normal distribution is still widely applicable in that, for many such cases, the physical limits are located far from the population mean.

The Uniform Distribution

The uniform distribution has limited applicability to errors in measurement and has virtually no applicability to biases of equipment parameters. For the latter, it can be argued that it is not a physically credible distribution. Despite its "unphysical" character, applying the uniform distribution to obtain Type B uncertainty estimates is a practice that has been gaining ground over the past few years. There are two main reasons for this:

Ease of Use

Applying the uniform distribution makes it easy to obtain an uncertainty estimate. If the limits $\pm a$ of the distribution are known, the uncertainty estimate is computed using Eq. (6).

It should be said that the "ease of use" advantage should be tempered by the following warning:[2]

> "When a component of uncertainty is determined in this manner contributes significantly to the uncertainty of a measurement result, it is prudent to obtain additional data for its further evaluation."

At our present level of analytical development,[3] the ease of use advantage is more difficult to justify than previously. Given the emergence of new methods and tools, the only "excuse" for not using a physically realistic distribution, applying a containment probability and estimating the degrees of freedom (if appropriate) is that the information for doing so is either (1) not at our fingertips or (2) it puts too much of a strain on technicians that are not trained in estimating uncertainty.[4]

GUM Authority

It has been asserted by some that the use of the uniform distribution is (uniformly?) recommended in the GUM.[5] This is not true. In fact, most of the methodology of the GUM is based on the assumption that the underlying error distribution is normal.

---

[2] Quoted from Section 4.3.7 of the GUM.

[3] Castrup, S., "A Comprehensive Comparison of Uncertainty Analysis Tools," *Proc. Measurement Science Conference*, January 2004, Anaheim.

[4] Since the goal of ensuring accurate measurements often involves the investment of considerable management and engineering time, equipment acquisition expenditures and proficiency training costs, it seems curious that, just when we get to the payoff, we become lazy and sloppy.

[5] ANSI/NCSL Z540-2, U.S. Guide to the Expression of Uncertainty in Measurement, 1st Ed., 1997, Boulder.

Another source of confusion is that some of the examples in the GUM apply the uniform distribution in situations that appear to be incompatible with its use. It is reasonable to suppose that much of this is due to the fact that rigorous Type B estimation methods and tools were not available at the time the GUM was published, and the uniform distribution was an "easy out." As stated above, the lack of such methods and tools has since been rectified. For clarification on this issue, the reader is referred to Section 4.3 of the GUM.

Uniform Distribution Alternatives

There are two types of uniform distribution. One describes errors that fall within symmetric limits ±*a* centered at zero. This type of uniform distribution is classified as a “round-off” distribution. It is shown in Figure 2.

The second type of uniform distribution describes errors that are distributed between the limits 0 and *a*. This distribution is classified as a “truncation” distribution.[6] It is shown in Figure 3.

The pdfs

Round-off Distribution

The round-off uniform distribution is defined by the probability density function

$$f(\varepsilon) = \begin{cases} \dfrac{1}{2a}, & -a \le \varepsilon \le a \\ 0, & \text{otherwise}. \end{cases} \tag{4}$$

Truncation Distribution

The truncation uniform distribution is defined by the probability density function

$$f(\varepsilon) = \begin{cases} \dfrac{1}{a}, & 0 \le \varepsilon \le a \\ 0, & \text{otherwise}. \end{cases} \tag{5}$$

The Variances

Round-off Distribution

By Eqs. (2) and (4), the variance is given by

$$\begin{aligned} \sigma^2 &= \int_{-a}^{a} f(\varepsilon)\varepsilon^2 d\varepsilon \\ &= \frac{1}{2a}\int_{-a}^{a} \varepsilon^2 d\varepsilon \\ &= \frac{1}{2a}\frac{\varepsilon^3}{3}\bigg|_{-a}^{a} = \frac{a^2}{3}. \end{aligned}$$

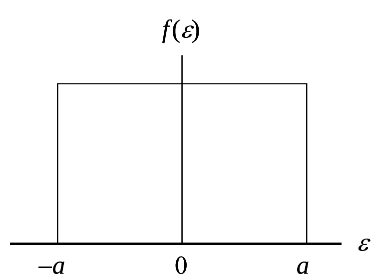

Figure 2. The Round-off Uniform Distribution. Shown is the pdf for a uniformly distributed random variable ε with mean zero and bounding limits ±*a*.

[6] To the best of the author’s knowledge, this distribution was first proposed in June 2008 by Paul Reese of Wyle Laboratories.

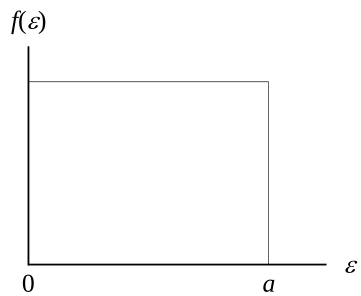

Figure 3. The Truncation Uniform Distribution. Shown is the pdf for a random variable ε uniformly distributed between 0 and a.

Truncation Distribution

By Eqs. (2) and (5), the variance is given by

$$\begin{aligned}\sigma^2 &= \int_0^a f(\varepsilon)\varepsilon^2 d\varepsilon \\ &= \frac{1}{a}\int_0^a \varepsilon^2 d\varepsilon \\ &= \frac{1}{a}\frac{\varepsilon^3}{3}\bigg|_0^a = \frac{a^2}{3}.\end{aligned}$$

The Standard Deviations

The standard deviation σ for both the round-of and truncation distributions is

$$\sigma = \frac{a}{\sqrt{3}}. \tag{6}$$

The Distribution Limits

For this expression to be useful, the limits of the distribution must be *minimum containment limits*. In practice, we may not know the value of a, and should attempt to obtain a containment probability p and containment limits $\pm L$ for the round-off distribution and an upper containment limit L for the truncation distribution, from whence we get

$$a = \frac{L}{p}, \quad L \le a . \tag{7}$$

The Triangular Distribution

The triangular distribution has been proposed by the authors of the GUM for use in cases where the containment probability is 100%, but there is a central tendency for values of the variable of interest, as shown in Figure 4. The triangular distribution is the simplest distribution possible with these characteristics.

The triangular distribution is also found in cases where two uniformly distributed errors with the same mean and bounding limits are combined linearly.

The pdf

The pdf for the distribution is

$$f(\varepsilon) = \begin{cases} (a+\varepsilon)/a^2, & -a \le \varepsilon \le 0 \\ (a-\varepsilon)/a^2, & 0 \le \varepsilon \le a \\ 0, & \text{otherwise.} \end{cases} \tag{8}$$

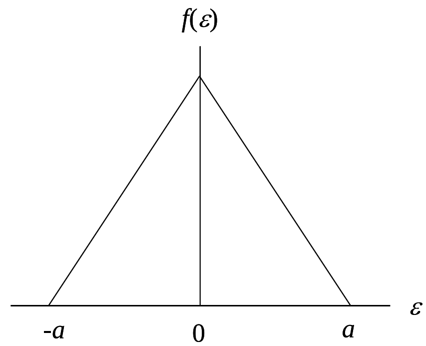

Figure 4. The Triangular Distribution. Shown is the pdf for a random variable ε with mean zero and bounding limits $\pm a$.

The Variance

By Eqs. (2) and (8), the variance is given by

$$\begin{aligned}
\sigma^2 &= \int_{-a}^{a} f(\varepsilon)\varepsilon^2 d\varepsilon \\
&= \frac{1}{a^2}\int_{-a}^{0}(a+\varepsilon)\varepsilon^2 d\varepsilon + \frac{1}{a^2}\int_{0}^{a}(a-\varepsilon)\varepsilon^2 d\varepsilon \\
&= \frac{2}{a^2}\int_{0}^{a}(a-\varepsilon)\varepsilon^2 d\varepsilon \\
&= \frac{2}{a^2}\left(a\frac{\varepsilon^3}{3} - \frac{\varepsilon^4}{4}\right)_0^a \\
&= 2a^2(1/3 - 1/4) = \frac{a^2}{6}.
\end{aligned}$$

The Standard Deviation

The standard deviation σ is the square root of the variance. Hence

$$\sigma = \frac{a}{\sqrt{6}} . \tag{9}$$

The Distribution Limits

In cases where a containment probability $p < 1$ can be determined for known containment limits $\pm L$, where $L < a$, the limits of the distribution are given by

$$a = \frac{L}{1-\sqrt{1-p}}, \quad L \le a . \tag{10}$$

Applicability of the Triangular Distribution

This distribution can be said to apply to linear combinations of pairs of uniformly distributed errors, as described earlier, and to post-test distributions under certain restricted conditions. It has also shown promise when applied to errors involved in interpolating between tabulated values. In these cases, the minimum error occurs at each tabulated value and the maximum error can often be assumed to occur at the mid point between these values.

Apart from these cases, the triangular distribution has limited applicability to physical errors or deviations. Like the uniform distribution, it displays abrupt transitions at the bounding limits and at the zero point, behaviors which are physically unrealistic for measurement errors. In addition, the linear slopes of the pdf are somewhat fanciful.

Other distributions, such as the quadratic, cosine and half-cosine, are available that are as easy to specify and use as the uniform or triangular and have none of their drawbacks.

The Quadratic Distribution

A distribution that eliminates the abrupt change at the zero point, does not exhibit unrealistic linear probability slopes and satisfies the need for a central tendency is the quadratic distribution.

The pdf

This distribution is defined by the pdf shown in Figure 5

$$f(\varepsilon) = \begin{cases} \frac{3}{4a}\left[1-(\varepsilon/a)^2\right], & -a \le \varepsilon \le a \; . \\ 0\,, & \text{otherwise} \end{cases} \tag{11}$$

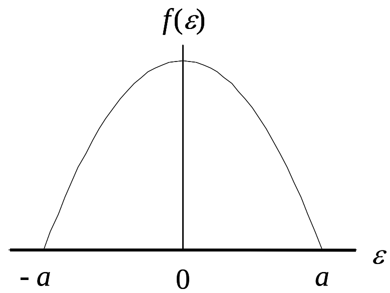

Figure 5. The Quadratic Distribution. Exhibits a central tendency without discontinuities and does not assume linear pdf behavior.

The Variance

By Eq. (2),

$$\begin{aligned} \sigma^2 &= \frac{3}{4a}\int_{-a}^{a}\left[1-(\varepsilon/a)^2\right]\varepsilon^2\,d\varepsilon \\ &= \frac{3}{4a}\left(\frac{\varepsilon^3}{3}-\frac{\varepsilon^5}{5a^2}\right)\Bigg|_{-a}^{a} \\ &= \frac{3a^2}{2}\left(\frac{1}{3}-\frac{1}{5}\right) = \frac{a^2}{5}. \end{aligned}$$

The Standard Deviation

The standard deviation for this distribution is the square root of the variance:

$$\sigma = \frac{a}{\sqrt{5}}\,. \tag{12}$$

The Distribution Limits

For a containment probability p and containment limits $\pm L$, the minimum bounding limits $\pm a$ are obtained from

$$a = \frac{L}{2p}\left(1+2\cos\left[\frac{1}{3}\arccos(1-2p^2)\right]\right) \quad -1 < p < 1\,. \tag{13}$$

This expression was developed by Mark Kuster of BWXT Pantex LLC. Its derivation is given in the Appendix to this monograph.

Applicability of the Quadratic Distribution

The quadratic distribution can be applied to error sources whose errors are known to have finite symmetric bounding limits and tend to not aggregate closely around the mid point of these limits. While it displays abrupt transitions at the bounding limits as do the uniform and triangular distributions, it is continuous at the zero point, and, therefore, is considered to be more physically realistic than either the uniform or the triangular.

The Cosine Distribution

While the quadratic distribution eliminates discontinuities within the bounding limits, it rises abruptly *at* the limits. Although the quadratic distribution has wider applicability than either the triangular or uniform distribution, this feature nevertheless diminishes its physical validity. A distribution that overcomes this shortcoming, exhibits a central tendency and can be determined from minimum containment limits is the cosine distribution, shown in Figure 6.

The pdf

The pdf for this distribution is given by

$$f(\varepsilon) = \begin{cases} \frac{1}{2a}\left[1 + \cos\left(\frac{\pi\varepsilon}{a}\right)\right], & -a \le \varepsilon \le a \ . \\ 0\,, & \text{otherwise} \end{cases} \tag{14}$$

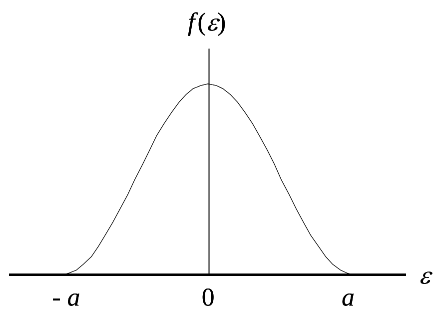

Figure 6. The Cosine Distribution. A 100% containment distribution with a central tendency and lacking discontinuities.

The Variance

Substituting Eq. (14) in Eq. (2) yields

$$\begin{aligned}
\sigma^2 &= \frac{1}{2a}\int_{-a}^{a}\left[1 + \cos\left(\frac{\pi\varepsilon}{a}\right)\right]\varepsilon^2 d\varepsilon \\
&= \frac{1}{2a}\left[\frac{\varepsilon^3}{3}\Bigg|_{-a}^{a} + \int_{-a}^{a}\varepsilon^2\cos\left(\frac{\pi\varepsilon}{a}\right)d\varepsilon\right] \\
&= \frac{a^2}{3} + \frac{a^2}{2\pi^3}\int_{-\pi}^{\pi}\zeta^2\cos\zeta\, d\zeta \\
&= \frac{a^2}{3} + \frac{a^2}{2\pi^3}\left[2\zeta\cos\zeta + (\zeta^2 - 2)\sin\zeta\right]\Bigg|_{-\pi}^{\pi} \\
&= \frac{a^2}{3} - \frac{2a^2}{\pi^2} = \frac{a^2}{3}\left(1 - \frac{6}{\pi^2}\right).
\end{aligned}$$

The Standard Deviation

The standard deviation is the square root of the variance:

$$\sigma = \frac{a}{\sqrt{3}}\sqrt{1 - \frac{6}{\pi^2}} \ . \qquad (15)$$

The Distribution Limits

Solving for a when a containment probability p and containment limits $\pm L$ are given requires applying numerical iterative method to the expression

$$\frac{a}{\pi}\sin(\pi L / a) - ap + L = 0, \quad L \le a \ .$$

The solution algorithm has been implemented in various ISG products and freeware applications. It yields, for the ith iteration,

$$a_i = a_{i-1} - F / F' \ ,$$

where

$$F = \frac{a}{\pi}\sin(\pi L / a) - ap + L$$

and

$$F' = \frac{1}{\pi}\sin(\pi L / a) - \frac{L}{a}\cos(\pi L / a) - p \ .$$

The Half-Cosine Distribution

The half-cosine distribution, shown in Figure 7, is used in cases where the central tendency is not as pronounced as when the normal or cosine distribution would be appropriate. In this regard, it resembles the quadratic distribution without the discontinuities at the distribution limits.

The pdf

The pdf is given by

$$f(\varepsilon) = \begin{cases} \frac{\pi}{4a}\cos\left(\frac{\pi\varepsilon}{2a}\right) \ , & a \le \varepsilon \le a \ . \\ 0 \ , \quad \text{otherwise} & \end{cases} \qquad (16)$$

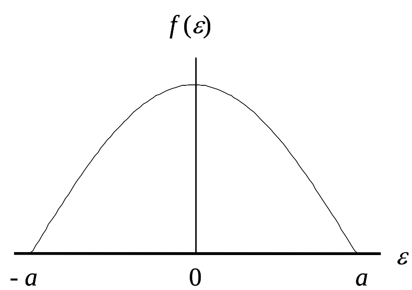

Figure 7. The Half-Cosine Distribution. Possesses a central tendency but exhibits a higher probability of occurrence near the minimum limiting values than either the cosine or the normal distribution.

The Variance

Using Eq. (16) in Eq. (2) gives

$$\begin{aligned}\sigma^2 &= \frac{\pi}{4a}\int_{-a}^{a}\cos\left(\frac{\pi\varepsilon}{2a}\right)\varepsilon^2 d\varepsilon \\ &= \frac{2a^2}{\pi^2}\int_{-\pi/2}^{\pi/2}\zeta^2\cos\zeta\, d\zeta \\ &= \frac{2a^2}{\pi^2}\left[2\zeta\cos\zeta + (\zeta^2-2)\sin\zeta\right]\Big|_{-\pi/2}^{\pi/2} \\ &= \frac{2a^2}{\pi^2}\left(\frac{\pi^2}{2}-4\right) = a^2\left(1-\frac{8}{\pi^2}\right).\end{aligned}$$

The Standard Deviation

The standard deviation is the square root of the variance:

$$\sigma = a\sqrt{1-\frac{8}{\pi^2}}\ . \tag{17}$$

The Distribution Limits

If containment limits $\pm L$ and a containment probability p are known, the limiting values may be obtained from the relation

$$a = \frac{\pi L}{2\sin^{-1}(p)},\quad L \le a\ . \tag{18}$$

Applicability of the Half-Cosine Distribution

As stated above, the half-cosine distribution can be applied to error sources whose errors are known to have finite symmetric bounding limits but do not display a tendency to aggregate closely around the mid point. It has the appearance of the quadratic distribution without the drawback of being discontinuous at the bounding limits.

Since the half-cosine distribution is continuous both throughout its range and at the bounding limits, it is more physically palatable than the uniform, triangular or quadratic distributions,.

The U Distribution

The following is a derivation of the U or "U-Shaped" distribution, shown in Figure 8, The U distribution is useful for specifying sinusoidal phase-varying subject parameters.

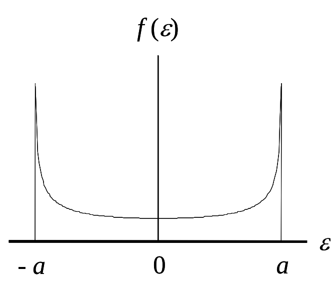

Figure 8. The U Distribution. The distribution is the pdf for sine waves of random phase incident on a plane.

The pdf

Preliminaries

The U distribution indicates the probability of sampling a particular sine wave amplitude from a signal whose phase is uniformly distributed. The relevant expressions are

$$\varepsilon = a \sin \varphi \ ,$$

where the phase angle distribution is uniform over [-π,π]

$$h(\varphi) = \begin{cases} 1/2\pi, & -\pi \le \varphi \le \pi \\ 0, & \textit{otherwise.} \end{cases}$$

Change of Variable

The pdf for ε can be obtained by using the change of variable technique:

$$f(\varepsilon) = \left| \frac{d\varphi}{d\varepsilon} \right| h(\varepsilon(\varphi)) \ ,$$

which yields

$$f(\varepsilon) = \begin{cases} \dfrac{1}{\pi\sqrt{a^2 - \varepsilon^2}}, & -a < x < a \\ 0, & \text{otherwise.} \end{cases} \tag{19}$$

The Variance

Substituting Eq. (19) in Eq. (2) gives

$$\begin{aligned} \sigma^2 &= \frac{1}{\pi} \int_{-a}^{a} \frac{\varepsilon^2}{\sqrt{a^2 - \varepsilon^2}} d\varepsilon \\ &= \frac{1}{\pi} \left(-\frac{\varepsilon\sqrt{a^2 - \varepsilon^2}}{2} + \frac{a^2}{2} \sin^{-1}(\varepsilon / a) \right) \Bigg|_{-a}^{a} \\ &= \frac{1}{\pi} \frac{a^2}{2} \pi = \frac{a^2}{2}. \end{aligned}$$

The Standard Deviation

The standard deviation is the square root of the variance:

$$\sigma = \frac{a}{\sqrt{2}} \ . \tag{20}$$

The Distribution Limits

If containment limits $\pm L$, where $L < a$ and a containment probability $p < 1$ are known, the parameter a can be estimated according to

$$a = \frac{L}{\sin\left(\pi p / 2\right)}, \quad L \le a \ . \tag{21}$$

Applicability of the U Distribution

The U distribution applies to errors that vary sinusoidally with time, such as RF signals incident on a load, or to quantities, such as temperatures or pressures maintained by automatic environmental control systems. Since the distribution is derived from the properties of such phenomena, it is a physically realistic distribution.

The Lognormal Distribution[7]

The lognormal distribution can often be used to estimate the uncertainty in equipment parameter bias in cases where the tolerance limits are asymmetric. It is also used in cases where a physical limit is present that lies close enough to

[7] For a more complete discussion see the article "The Lognormal Distribution," at www.isgmax.com.

the nominal or mode value to skew the error pdf in such a way that the normal distribution is not applicable.

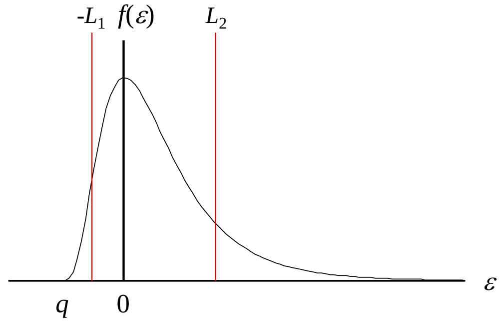

Figure 9. The Lognormal Distribution. Useful for describing distributions for parameters constrained by a physical limit or possessing asymmetric tolerances. The case shown is a "right-handed" distribution in which the mode value (zero) is greater than the physical limit *q*.

The pdf

The pdf for a lognormally distributed error ε is given by

$$f(\varepsilon) = \frac{1}{\sqrt{2\pi}\lambda\,|\,\varepsilon - q\,|} \exp\left\{ -\left[\ln\left(\frac{\varepsilon - q}{m - q} \right) \right]^2 \Big/ 2\lambda^2 \right\}, \tag{22}$$

where q is a physical limit for ε, m is the distribution median and λ is referred to as the "shape parameter." Figure 9 shows an example of a "right-handed" lognormal for which $q < 0$.

The Mean Value

The mean value for the distribution of interest is obtained by integrating the variable ε in Eq. (2) from q to ∞. Using Eq. (22), we have

$$\langle\varepsilon\rangle = \frac{1}{\sqrt{2\pi}\lambda} \int_q^\infty \exp\left\{ -\left[\ln\left(\frac{\varepsilon - q}{m - q} \right) \right]^2 \Big/ 2\lambda^2 \right\} \frac{\varepsilon}{(\varepsilon - q)} d\varepsilon$$

We define the variable ζ as

$$\zeta = \frac{\varepsilon - q}{m - q}.$$

Then

$$\varepsilon = q + (m - q)\zeta\ ,$$

and

$$d\varepsilon = (m - q)d\zeta\ .$$

With these substitutions, the mean value equation becomes

$$\begin{aligned} \langle\varepsilon\rangle &= \frac{1}{\sqrt{2\pi}\lambda} \int_0^\infty \exp\left\{ -\frac{[\ln(\zeta)]^2}{2\lambda^2} \right\} [q + (m - q)\zeta] \frac{d\zeta}{\zeta} \\ &= I_1 + I_2\,, \end{aligned} \tag{23}$$

where

$$I_1 = \frac{q}{\sqrt{2\pi}\lambda}\int_0^\infty \exp\left\{-\frac{[\ln(\zeta)]^2}{2\lambda^2}\right\}\frac{d\zeta}{\zeta}$$
$$= q$$

and

$$I_2 = \frac{\mu - q}{\sqrt{2\pi}\lambda}\int_0^\infty \exp\left\{-\frac{(\ln\zeta)^2}{2\lambda^2}\right\}d\zeta$$
$$= (m-q)e^{\lambda^2/2}\,.$$

Combining these results in Eq. (23) gives

$$\langle\varepsilon\rangle = q + (m-q)e^{\lambda^2/2}\,. \tag{24}$$

Variance

The variance is given by

$$\begin{aligned}\sigma^2 &= \int_q^\infty \left(\varepsilon - \langle\varepsilon\rangle\right)^2 f(\varepsilon)d\varepsilon \\ &= \int_q^\infty \varepsilon^2 f(\varepsilon)d\varepsilon - \langle\varepsilon\rangle^2 \\ &= \langle\varepsilon^2\rangle - \langle\varepsilon\rangle^2.\end{aligned} \tag{25}$$

From Eq. (23), we have

$$\begin{aligned}\langle\varepsilon\rangle^2 &= q^2 + 2q(m-q)e^{\lambda^2/2} + (m-q)^2 e^{\lambda^2} \\ &= [q + (m-q)e^{\lambda^2/2}]^2\,.\end{aligned} \tag{26}$$

We now determine the mean of ε^2. Using Eq. (2) we have

$$\langle\varepsilon^2\rangle = \frac{1}{\sqrt{2\pi}\lambda}\int_q^\infty \exp\left\{-\left[\ln\left(\frac{\varepsilon-q}{m-q}\right)\right]^2 \Big/ 2\lambda^2\right\}\frac{\varepsilon^2}{\varepsilon-q}d\varepsilon\,. \tag{27}$$

We again define a variable ζ as

$$\zeta = \frac{\varepsilon - q}{m-q}\,.$$

Then

$$\varepsilon = q + (m-q)\zeta\,,$$

$$\varepsilon^2 = q^2 + 2q(m-q)\zeta + (m-q)^2\zeta^2\,,$$

and

$$d\varepsilon = (m-q)d\zeta\,.$$

With these substitutions, Eq. (26) becomes

$$\begin{aligned}\langle\varepsilon^2\rangle &= \frac{1}{\sqrt{2\pi}\lambda}\int_0^\infty \exp\left\{-\frac{(\ln\zeta)^2}{2\lambda^2}\right\}[q^2 + 2q(m-q)\zeta + (m-q)^2\zeta^2]\frac{d\zeta}{\zeta} \\ &= I_3 + I_4 + I_5\,,\end{aligned} \tag{28}$$

where

$$I_3 = \frac{q^2}{\sqrt{2\pi}\lambda}\int_0^\infty \exp\left\{-\frac{\left[\ln(\zeta)\right]^2}{2\sigma^2}\right\}\frac{d\zeta}{\zeta} \qquad (29)$$
$$= q^2,$$

$$I_4 = \frac{2q(m-q)}{\sqrt{2\pi}\lambda}\int_0^\infty \exp\left\{-\frac{\left[\ln(\zeta)\right]^2}{2\lambda^2}\right\}d\zeta$$
$$= \frac{2q(m-q)}{\sqrt{2\pi}}\int_{-\infty}^\infty e^{-\xi^2/2}e^{\lambda\xi}d\xi \qquad (30)$$
$$= 2q(m-q)e^{\lambda^2/2},$$

and

$$I_5 = \frac{(m-q)^2}{\sqrt{2\pi}\lambda}\int_0^\infty \exp\left\{-\frac{\left(\ln\zeta\right)^2}{2\lambda^2}\right\}\zeta d\zeta$$
$$= \frac{(m-q)^2}{\sqrt{2\pi}}\int_{-\infty}^\infty e^{-\xi^2/2}e^{2\lambda\xi}d\xi \qquad (31)$$
$$= (m-q)^2 e^{2\lambda^2}.$$

Combining Eqs. (29 – 31) in Eq. (28) gives

$$\begin{aligned}\langle\varepsilon^2\rangle &= q^2 + 2q(m-q)e^{\lambda^2/2} + (m-q)^2 e^{2\lambda^2} \\ &= [q + (m-q)e^{\lambda^2/2}]^2 + (m-q)^2 e^{\lambda^2}(e^{\lambda^2}-1).\end{aligned} \qquad (32)$$

Then, using Eqs. (32) and (26) in Eq. (25) yields

$$\begin{aligned}\sigma^2 &= \langle\varepsilon^2\rangle - \langle\varepsilon\rangle^2 \\ &= (m-q)^2 e^{\lambda^2}(e^{\lambda^2}-1).\end{aligned} \qquad (33)$$

Standard Deviation

The standard deviation is just the square root of the variance:

$$\sigma = (m-q)e^{\lambda^2/2}\sqrt{e^{\lambda^2}-1}\,. \qquad (34)$$

Other Statistics

In general, for a lognormally distributed error, we have the following relations:

Mode	0
Mean	$q + m(\mu - q)e^{\lambda^2/2}$
Median	$m = q + (\mu - q)e^{\lambda^2}$
Variance	$(m-q)^2 e^{\lambda^2}(e^{\lambda^2}-1)$
Standard Deviation	$\lvert m-q\rvert e^{\lambda^2/2}\sqrt{e^{\lambda^2}-1}$

The quantities m, q and λ are obtained by numerical iteration, given containment limits and containment probabilities. To date, the only known published analytical metrology applications that perform this process are

UncertaintyAnalyzer,[8] AccuracyRatio.[9]

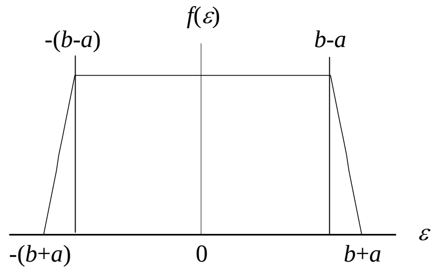

Figure 10. Trapezoidal Distribution. Shown is a distribution for the sum of two uniformly distributed variables with bounding values a = 3 and b = 4. The 100% containment limits are ± b and the uniform density limits are ±a.

The Trapezoidal Distribution

If two errors ε_x and ε_y are uniformly distributed with bounding values ±a and ±b, where $b \geq a$, then their sum

$$\varepsilon = \varepsilon_x + \varepsilon_y$$

follows a trapezoidal distribution with discontinuities at ±(b – a) and ±(b + a), as shown in Figure 10.

The pdf

Working with the pdf can be facilitated by making the transformations

$$c = b - a \tag{35}$$

and

$$d = b + a. \tag{36}$$

Then the pdf is given by

$$f(\varepsilon) = \begin{cases} \dfrac{1}{d^2 - c^2}(d + \varepsilon), & -d \leq \varepsilon \leq -c \\ \dfrac{1}{d + c}, & -c \leq \varepsilon \leq c \\ \dfrac{1}{d^2 - c^2}(d - \varepsilon), & c \leq \varepsilon \leq d \\ 0. & \text{otherwise} \end{cases} \tag{37}$$

The Variance

Averaging ε^2 over the bounding limits of this distribution yields the variance

[8] UncertaintyAnalyzer 3.0, © 2005-2008, Integrated Sciences Group, All Rights Reserved.

[9] AccuracyRatio 1.6, © 2004-2008, Integrated Sciences Group, All Rights Reserved.

$$
\begin{aligned}
\sigma^2 &= \frac{1}{d^2-c^2}\int_{-d}^{-c}\varepsilon^2(d+\varepsilon)d\varepsilon + \frac{1}{d+c}\int_{-c}^{c}\varepsilon^2 d\varepsilon + \frac{1}{d^2-c^2}\int_{c}^{d}\varepsilon^2(d-\varepsilon)d\varepsilon \\
&= \frac{2}{d^2-c^2}\int_{c}^{d}\varepsilon^2(d-\varepsilon)d\varepsilon + \frac{2}{d+c}\int_{0}^{c}\varepsilon^2 d\varepsilon \\
&= \frac{2}{d^2-c^2}\left[\frac{1}{3}d(d^3-c^3) - \frac{1}{4}(d^4-c^4)\right] + \frac{2}{3}\frac{c^3}{d+c} \\
&= \frac{d^4-c^4}{6(c^2-c^2)} = \frac{d^2+c^2}{6}.
\end{aligned}
\tag{38}
$$

Substituting from Eqs. (35) and (36) in Eq. (38) gives

$$
\begin{aligned}
\sigma^2 &= \frac{(b+a)^2+(b-a)^2}{6} \\
&= \frac{a^2+b^2}{3}.
\end{aligned}
\tag{39}
$$

The Standard Deviation

From Eq. (39), the standard deviation is

$$
\sigma = \sqrt{\frac{a^2+b^2}{3}}\ . \tag{40}
$$

Applicability of the Trapezoidal Distribution

The trapezoidal applies to the sum of two uniformly distributed errors. Apart from this fact, it is difficult to imagine an instance where it would be applicable on it own merit. It has been recommended in cases where 100% containment limits are known, the probability density is believed to be less near the limits than at their midpoint and there exists a region inside the limits where the pdf is approximately uniform.

Given these considerations, we might be more inclined to use the quadratic, cosine, half-cosine or utility distribution (see below) on the grounds that each exhibits a more physically realistic character. Another contrary point in its disfavor is that the construction of the trapezoidal distribution requires not only the specification of 100% containment limits but uniform density limits as well. In the author's experience, obtaining the limits of the trapezoidal distribution from technical expertise has not been a fruitful exercise.

The Utility Distribution

One of the key variables in evaluating the return on investment of alternative technical decisions or policies is a quantity called *utility*. A utility function that has been applied to evaluating ROI for test and calibration support hierarchies[10] is one whose behavior is somewhat similar to that of the trapezoidal distribution. Specifically, this function has a region of approximately uniform utility, bounded by limits that mark points where the utility begins to decrease from its maximum value, eventually reaching limits that correspond to zero utility.

The pdf

The pdf for the utility distribution shown in Figure 11 is

[10] Castrup, H., "Calibration Requirements Analysis System," *Proc. NCSL Workshop & Symposium*, Denver, July 9-13, 1989.

$$f(\varepsilon)=\begin{cases}\dfrac{1}{a+b}, & |\varepsilon|\le a\\ \dfrac{1}{a+b}\cos^2\left[\dfrac{\pi(|\varepsilon|-a)}{2(b-a)}\right], & a\le|\varepsilon|\le b\\ 0, & |\varepsilon|\ge b.\end{cases} \tag{41}$$

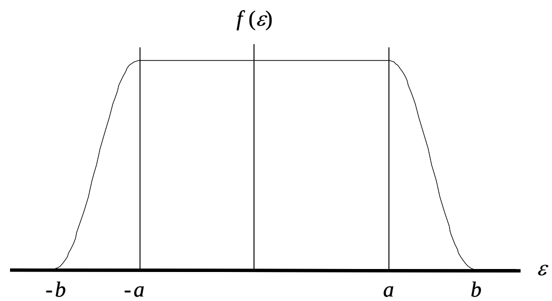

Figure 11. The Utility Distribution. The probability of occurrence is approximately uniform between the values ±*a*, tapering off to zero at the limits ±*b*.

The Variance

Integrating ε^2 over the bounding limits of this distribution yields the variance

$$\begin{aligned}\sigma^2&=\frac{1}{a+b}\int_{-b}^{-a}\cos^2\left[\frac{\pi(\varepsilon+a)}{2(b-a)}\right]\varepsilon^2d\varepsilon+\frac{1}{a+b}\int_{-a}^{a}\varepsilon^2d\varepsilon+\frac{1}{a+b}\int_{a}^{b}\cos^2\left[\frac{\pi(\varepsilon-a)}{2(b-a)}\right]\varepsilon^2d\varepsilon\\&=\frac{2}{a+b}\int_{a}^{b}\cos^2\left[\frac{\pi(\varepsilon-a)}{2(b-a)}\right]\varepsilon^2d\varepsilon+\frac{2}{a+b}\int_{0}^{a}\varepsilon^2d\varepsilon\\&=\frac{4(b-a)}{\pi(a+b)}\int_{0}^{\pi/2}\cos^2\zeta\left[\frac{2}{\pi}(b-a)\zeta+a\right]^2d\zeta+\frac{2a^3}{3(a+b)}\\&=\frac{4(b-a)}{\pi(a+b)}\int_{0}^{\pi/2}\left[\frac{4}{\pi^2}(b-a)^2\zeta^2+\frac{4}{\pi}a(b-a)\zeta+a^2\right]\cos^2\zeta\,d\zeta+\frac{2a^3}{3(a+b)}\\&=I_1+I_2+I_3+\frac{2a^3}{3(a+b)},\end{aligned} \tag{42}$$

where

$$\begin{aligned}I_1&=\frac{16(b-a)^3}{\pi^3(a+b)}\int_{0}^{\pi/2}\zeta^2\cos^2\zeta\,d\zeta\\&=\frac{16(b-a)^3}{\pi^3(a+b)}\left(\frac{\zeta^3}{6}+\frac{1}{4}\zeta^2\sin2\zeta-\frac{1}{8}\sin2\zeta+\frac{1}{4}\zeta\cos2\zeta\right)_0^{\pi/2}\\&=\frac{2(b-a)^3}{\pi^2(a+b)}\left(\frac{\pi^2}{6}-1\right),\end{aligned} \tag{43}$$

$$\begin{aligned} I_2 &= \frac{4(b-a)}{\pi(a+b)}\frac{4}{\pi}a(b-a)\int_0^{\pi/2}\zeta\cos^2\zeta\,d\zeta \\ &= \frac{16a(b-a)^2}{\pi^2(a+b)}\int_0^{\pi/2}\zeta\cos^2\zeta\,d\zeta \\ &= \frac{16a(b-a)^2}{\pi^2(a+b)}\left(\frac{\zeta^2}{4}+\frac{\zeta\sin 2\zeta}{4}+\frac{\cos 2\zeta}{8}\right)_0^{\pi/2} \\ &= \frac{a(b-a)^2}{\pi^2(a+b)}\left(\pi^2-4\right), \end{aligned} \tag{44}$$

and

$$\begin{aligned} I_3 &= \frac{4a^2(b-a)}{\pi(a+b)}\int_0^{\pi/2}\cos^2\zeta\,d\zeta \\ &= \frac{4a^2(b-a)}{\pi(a+b)}\left(\frac{\zeta}{2}+\frac{\sin 2\zeta}{4}\right)_0^{\pi/2} \\ &= \frac{a^2(b-a)}{a+b}. \end{aligned} \tag{45}$$

Substituting Eqs. (43) – (45) in Eq. (42) gives

$$\begin{aligned} \sigma^2 &= I_1+I_2+I_3+\frac{2a^3}{3(a+b)} \\ &= \frac{2(b-a)^3}{\pi^2(a+b)}\left(\frac{\pi^2}{6}-1\right)+\frac{a(b-a)^2}{\pi^2(a+b)}\left(\pi^2-4\right)+\frac{a^2(b-a)}{a+b}+\frac{2a^3}{3(a+b)} \\ &= \frac{1}{a+b}\left[2(b-a)^3\left(\frac{1}{6}-\frac{1}{\pi^2}\right)+a(b-a)^2\left(1-\frac{4}{\pi^2}\right)+a^2b-\frac{1}{3}a^3\right] \\ &= \frac{1}{a+b}\left[2(b^3-3ab^2+3a^2b-a^3)\left(\frac{1}{6}-\frac{1}{\pi^2}\right)+a(b^2-2ab+a^2)\left(1-\frac{4}{\pi^2}\right)+a^2b-\frac{1}{3}a^3\right] \\ &= \frac{1}{a+b}\left[\frac{b^3+a^3}{3}-\frac{2}{\pi^2}\left(b^3-ab^2-a^2b+a^3\right)\right] \\ &= \frac{b^3+a^3}{3(b+a)}-\frac{2}{\pi^2}(b-a)^2. \end{aligned}$$

The Standard Deviation

Taking the square root of the variance gives

$$\sigma = \sqrt{\frac{b^3+a^3}{3(b+a)}-\frac{2}{\pi^2}(b-a)^2}\ . \tag{46}$$

Applicability of the Utility Distribution

The character of this function makes it arguably more useful than the trapezoidal distribution in that it is free of discontinuities and requires the same information for its specification. While obtaining the necessary information for constructing a utility function is fairly straightforward, obtaining the limits ± a for the utility distribution is hampered by the same practical difficulties encountered in obtaining the limits for the uniform density portion of the trapezoidal distribution.

Recommendations for Selecting an Error Source Distribution

The following are offered as guidelines for selecting an appropriate error distribution:

1. Unless information to the contrary is available, the normal distribution should be applied as the default distribution.

2. If it is suspected that the distribution of the value of interest is skewed, apply the lognormal distribution.

   In using the normal or lognormal distribution, some effort must be made to estimate a containment probability. If a set of containment limits is available, but 100% containment has been observed, then the following is recommended:

3. If the value of interest has been subjected to random usage or handling stress, and is assumed to possess a central tendency, apply the cosine distribution. If it is suspected that values are more evenly distributed, apply either the quadratic or half-cosine distribution, as appropriate. The triangular distribution may be applicable to estimating uncertainty due to interpolation errors, and, under certain rare circumstances, when dealing with parameters following testing or calibration.

4. If the amplitude of the value of interest varies sinusoidally with time, apply the U distribution.

5. If the value of interest is the resolution uncertainty of a digital readout, apply the uniform distribution. This distribution is also applicable to estimating the uncertainty due to quantization error and the uncertainty in RF phase angle.