An Investigation into Estimating Type B Degrees of Freedom

H. Castrup
President, Integrated Sciences Group
June 12, 2001

Background

The degrees of freedom associated with an uncertainty estimate quantifies the amount of information that went into the estimate. As this amount of information grows, the degrees of freedom becomes larger. Conversely, if our information is meager or unclear, the degrees of freedom shrinks. Another way of saying this is that, if the uncertainty or variance in an uncertainty estimate is large, the degrees of freedom is small. An approximate expression of this relationship, as applied to computing the degrees of freedom for a Type B estimate, is provided in Eq. G.3 of the GUM [1]

$$\nu \simeq \frac{1}{2}\frac{u^2(x)}{\sigma^2[u(x)]} . \qquad (1)$$

This is an intuitively appealing result in that the degrees of freedom is inversely proportional to the square of the uncertainty in the uncertainty estimate. Hence, a large variance in the uncertainty estimate yields a small degrees of freedom.

While the relation is appealing, the GUM offers no advice on how to compute $\sigma[u(x)]$. Fortunately, a method was developed since the publication of the GUM, that was reported at the 2000 MSC [2], and has proved useful in practice. This method will be referred to herein as the Castrup method.

Recent Findings

An investigation of has been undertaken [3] in which sets of normal deviates are generated with known standard deviations for a variety of sample sizes. Following this, standard deviations for each set are calculated using both Type A and Type B procedures.

The Type A estimate is, of course,

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

while the Type B estimate is given by

$$u_x = \frac{L}{\Phi^{-1}\left(\frac{1+p}{2}\right)} . \qquad (3)$$

In this expression, the variable L is an arbitrary limit used to bound values of x around the mean value and p is the observed fraction of values found within $\pm L$ of the mean. For example, if u_x is the uncertainty in the bias of a toleranced parameter, $\pm L$ would most likely be the tolerance limits for a parameter's deviation from nominal and p would be the observed fraction in-tolerance. In discussing Type B uncertainty estimates, the limits $\pm L$ are referred to as containment limits and p is called the *containment probability*.

Preliminary Results

The results to date show that, for most of the simulated trials, as the sample size n grows to 50 or more, the Type B uncertainty estimate actually comes closer to the underlying population standard deviation (which is a known parameter in generating the random normal deviates) than does the Type A estimate.

However, using the Castrup method, the estimated Type B degrees of freedom turns out to be considerably less than the sample degrees of freedom for the Type A estimate. This is somewhat disquieting, since the Type B estimate is at least as good as the Type A estimate. Of course, obtaining a smaller degrees of freedom for the Type B estimate is a consequence of the difference in the mathematical forms of Eqs. (2) and (3).

Investigation

The question is, why doesn't the Castrup method yield a larger degrees of freedom? It may be that the method used to determine $\sigma[u(x)]$ is flawed, or the GUM relation is flawed, or both. As a first step in isolating the problem, an attempt was made to independently derive Eq. (1). The attempt begins with a consideration of the variance of a sample standard deviation.

Let s_ν represent the standard deviation, taken on a sample of size $n = \nu + 1$ of a N(0, u^2) variable x. We know that the quantity $\nu s_\nu^2 / u^2$ is χ^2-distributed with ν degrees of freedom. The χ^2-distribution has the pdf

$$f(x) = \frac{x^{(\nu-1)/2} e^{-x/2}}{2^{\nu/2}\Gamma\left(\frac{\nu}{2}\right)} .$$

Accordingly, we set $x = \nu s_\nu^2 / u^2$, or $s_\nu^2 = \left(u^2/\nu\right)x$, and compute the variance in s_ν^2.

$$\sigma^2\left(s_\nu^2\right) = \mathrm{var}\left(s_\nu^2\right) = \frac{u^4}{\nu^2}\mathrm{var}\left(x\right) .$$

For a χ^2-distributed x, we have

$$\mathrm{var}\left(x\right) = 2\nu$$

Substituting in the variance of s_ν^2 yields

$$\nu = 2\frac{u^4}{\sigma^2\left(s_\nu^2\right)} , \tag{4}$$

as contrasted with Eq. G3 of the GUM

$$\nu_{GUM} \simeq \frac{1}{2}\frac{u^2}{\sigma^2\left(s_\nu\right)} . \tag{5}$$

We will now compare these alternative expressions by using the Castrup method to obtain $\sigma^2\left(s_\nu\right)/u^2$ and $\sigma^2\left(s_\nu^2\right)/u^4$. Using Eq. (13) of Ref 2, we have

$$\frac{\sigma^2\left(s_\nu\right)}{u^2} \simeq \frac{u_L^2}{L^2} + \frac{1}{\varphi^2}\frac{\pi}{2}e^{\varphi^2}u_p^2 , \tag{6}$$

where u_L is the uncertainty in the containment limit L and u_p is the uncertainty in the containment probability p. the function φ is defined as

$$\varphi(p) = \Phi^{-1}\left[(1+p)/2\right] .$$

Using the method of Ref 2, we can show that

$$\frac{\sigma^2\left(s_\nu^2\right)}{u^4} \simeq 4\left(\frac{u_L^2}{L^2} + \frac{\pi}{2}e^{\varphi^2}\frac{u_p^2}{\varphi^2}\right) . \tag{7}$$

Comparison of Eq. (6) with Eq. (7) reveals that

$$\frac{\sigma^2\left(s_\nu^2\right)}{u^4} = 4\frac{\sigma^2(s_\nu)}{u^2} ,$$

which yields, with the aid of Eqs. (4) and (5),

$$\nu = 2\frac{u^4}{\sigma^2\left(s_\nu^2\right)} \simeq 2\left[\frac{u^2}{4\sigma^2\left(s_\nu\right)}\right] = \frac{1}{2}\frac{u^2}{\sigma^2\left(s_\nu\right)} = \nu_{GUM} . \tag{8}$$

Discussion

From the above, we can write

$$\sigma^2\left(s_\nu^2\right) \simeq 4u^2\sigma^2\left(s_\nu\right) . \tag{9}$$

However, from the properties of the χ^2 distribution, we have

$$\sigma^2\left(s_\nu^2\right) = \frac{2u^4}{\nu} ,$$

and

$$\sigma^2\left(s_\nu\right) = u^2\left(1 - \frac{2c_\nu^2}{\nu}\right),$$

where

$$c_\nu = \frac{\Gamma\left(\frac{\nu+1}{2}\right)}{\Gamma\left(\frac{\nu}{2}\right)} .$$

If Eq. (9) is valid, we have the relation

$$\frac{\Gamma\left(\frac{\nu+1}{2}\right)}{\Gamma\left(\frac{\nu}{2}\right)} \simeq \frac{\sqrt{2\nu-1}}{2} . \tag{10}$$

This relation is a consequence of Eq. (9), which is, in turn, a consequence of using the Castrup method. Also, by Eq. (8), the validity of ν_{GUM} appears to be dependent on the validity of the Castrup method, which also leads to Eq. (10).

I have not found the equality in Eq. (10) in any text, handbook or other reference. However, I have tested it by numerical calculation for values of ν ranging from 1 to 1000. It turns out that the equality is approximate for small values of ν (less than 50 or so) and improves as ν increases. For instance, it is exact to four decimal places at $\nu = 74$.

The approximate nature of the equality may be explained in that it springs from the Castrup method, which utilizes an expansion of the error in σ to first order in L and p, i.e., higher order terms are neglected. The approximate nature of the equality also indicates that the expression for ν in Eq. (5) is only approximately equivalent to the expression for ν in Eq. (4).

Utility of the Estimates

It is interesting to consider the case where $n = 100$. Suppose $\sigma = 1$ and the tolerance limits are ± 1.96. In this case, we would observe about 95% of the "observations" to be within the limits. This is what occurs when we simulate normal deviates. For example, a typical simulation using these parameters yielded 93% of observations within the limits ± 1.96.

The sample standard deviation was estimated using Eqs. (2) and (3) and the degrees of freedom were calculated using Eq. (8). The results are shown below.[1]

Table 1
A typical simulation with n = 100, σ = 1 and L = 1.96.

Description	Number of Observations	Number Observed within $\pm L$	Containment Probability p	Standard Deviation	Deg. Freedom
Population Value	-	-	-	1.0	∞
Type A Estimate	100	-	-	1.089	99
Type B Estimate	100	93	0.93	1.082	60

As has been stated, this example is typical of those for n = 100. It is interesting to note that, for such cases, the Type B standard deviation estimate is at least as good as the Type A estimate. Again, however, we have the result that the Type B degrees of freedom is smaller than the Type A value.

There are at least two possible explanations for this "information loss" that come to mind. One stems from the fact that the variance $\sigma^2\left(s_\nu^2\right)$ is evaluated using a first-order expansion of the error in s_ν^2, with higher-order terms neglected. Another possible explanation is that some loss of information may be due to using the binomial variance $p(1-p)/n$ to calculate u_p^2. If there is no uncertainty in L, this leads to the expression

$$\nu \simeq \frac{\varphi^2 e^{-\varphi^2}}{\pi p(1-p)} n ,$$

which will always be smaller than $n-1$. Perhaps a different approach to estimating u_p is in order.

Of course, it might be argued that a binomial statistic p, based on n independent Bernoulli trials, in which a value of 1 is assigned to a result within $\pm L$ and a value of 0 is assigned to a result external to $\pm L$, does not *really* contain as much information as a sample of data of size n. For instance, if we consider cases where p is nearly 1 or zero, i.e., cases where we may not obtain any values within $\pm L$ or external to $\pm L$, then we have little or no information from which to compute a Type B standard deviation. Given this argument, the smaller Type B degrees of freedom associated with computing u with Eq. (3) is an entirely satisfactory outcome.

The Castrup method produces estimates that are consistent with this observation. For example, ten simulations were carried out with n = 300, L = 1.96 and σ = 0.75. The small σ, relative to L, produced high values of p, resulting in poor estimates for σ in some of the trials. The Type B degrees of freedom ranged from 45 to 118. The results are shown in the table below. Note that, although different Type A uncertainty estimates were obtained for each simulation, the Type B uncertainty estimates tended to settle into groups. This is because the details of the sample (i.e., the sampled values) are "open" for the Type A estimates and masked or hidden for the Type B estimates. This supports the notion that a lesser amount of information is available for Type B estimates than for Type A estimates due to the binomial character of the former. Again, the smaller degrees of freedom for Type B estimates is to be expected.

[1] Experimenting with larger sample sizes and population standard deviations yield results that are consistent with those shown in the table.

Table 2
The results of ten simulations with n = 300, σ = 0.75 and L = 1.96.

Trial	**Type A Standard Deviation**	**Type B Standard Deviation**	**Containment Probability *p***	**Deg. Freedom**
1	0.7846	0.8425	0.9800	118
2	0.7531	0.7224	0.9933	108
3	0.7679	0.7609	0.9900	84
4	0.7343	0.7609	0.9900	84
5	0.7805	0.7609	0.9900	84
6	0.8000	0.8190	0.9833	108
7	0.7997	0.7920	0.9867	97
8	0.7674	0.7920	0.9867	97
9	0.7514	0.6677	0.9967	45
10	0.7708	0.7920	0.9867	97

References

[1] ANSI/NCSL, Z540-2-1997, *U.S. Guide to the Expression of Uncertainty in Measurement*, October 9, 1997.

[2] Castrup, H., "Estimating Category B Degrees of Freedom," Proc. Measurement Science Conference, Anaheim, January 2000.

[3] Work in-progress, Integrated Sciences Group, Bakersfield, CA 93306.