How to Gauge Unknown Unknowns: A Possible Theoretical Explanation of the Usual Safety Factor of 2

Joe Lorkowski and Vladik Kreinovich
Department of Computer Science
University of Texas at El Paso
500 W. University
El Paso, TX 79968, USA
lorkowski@computer.org, vladik@utep.edu

Abstract

To gauge the accuracy of a measuring instrument, engineers analyze possible factors contributing to the instrument's inaccuracy. In addition to *known* factors, however, there are usually *unknown* factors which also contribute to the instrument's inaccuracy. To properly gauge the instrument's accuracy – and thus, to make sure that we do not compromise our safety by underestimating the inaccuracy – we need to also take these "unknown unknowns" into account. In practice, this is usually done by multiplying the original estimate for inaccuracy by a "safety" factor of 2. In this paper, we provide a possible theoretical explanation for this empirical factor.

1 Safety Factor: Description of the Problem

What is a safety factor? When engineers design an accurate measuring instrument, they try their best to make it as accurate as possible. For this purpose, they apply known techniques to eliminate (or at least to drastically decrease) all major sources of error. For example:

- thermal noise can be drastically decreased by cooling,
- the effect of the outside electromagnetic fields can be drastically decreased by barriers made of conductive or magnetic materials,
- the effect of vibration can be decreased by an appropriate suspension, etc.

Once the largest sources of inaccuracy are (largely) eliminated, we need to deal with the next largest sources, etc. No matter how many sources of inaccuracy we deal with, there is always the next one. As a result, we can never fully eliminate all possible sources of inaccuracy.

At each stage of the instrument design, we usually understand reasonably well what is the main source of the remaining inaccuracy, and how to gauge the corresponding inaccuracy. As a result, we have a good estimate for the largest possible value Δ of the inaccuracy caused by this source.

This value Δ, however, does not provide a full description of the measurement inaccuracy. Indeed, in addition to the "known unknown" (the known main source of remaining inaccuracy), there are also "unknown unknowns" – smaller additional factors that also contribute to measurement inaccuracy.

To take these "unknown unknowns" into account, practitioners usually multiply Δ by a larger-than-one factor. This factor is known as the *safety factor*.

A related idea is used when we design engineering objects such as roads, bridges, houses, etc. For example, when we design a bridge, we want to make sure that its deformation caused by different loads does not exceed the limit after which catastrophic changes may occur. For this purpose, we first estimate the deformation caused by the known factors, and then – to be on the safe side – multiply this deformation by a safety factor.

Safety factor of 2 is most frequently used. In civil engineering, a usual recommendation is to use the safety factor of 2; see, e.g., [6].

This value has been in used for more than 150 years: according to [5], the earliest recorded use of this value can traced to a book published in 1858 [3].

This value is also widely used beyond standard civil engineering projects: e.g., it was used in the design of Buran, a successful Soviet-made fully automatic pilotless Space Shuttle [2].

Comment. It should be mentioned that in situations requiring extreme caution – e.g., in piloted space flights – usually, a larger value of safety factor is used, to provide additional safety.

Open problem. The fact that the safety factor of 2 has been in use for more than 150 years shows that this value is reasonable. However, there seems to be no convincing explanation of why this particular value is empirically reasonable.

In this paper, we provide a possible theoretical explanation for this value.

2 Safety Factor of 2: Our Explanation

Analysis of the problem. We know that in addition to the largest inaccuracy of size Δ, there is also next largest inaccuracy of size $\Delta_1 < \Delta$. Once we take that inaccuracy into account, then we will need to take into account the next inaccuracy, of size $\Delta_2 < \Delta_1$, etc. The final inaccuracy can be estimated as the sum of all these inaccuracies, i.e., as the sum $\Delta + \Delta_1 + \Delta_2 + \ldots$, where

$$\ldots < \Delta_{k+1} < \Delta_k < \ldots < \Delta_2 < \Delta_1 < \Delta.$$

To estimate this sum, it is therefore necessary to estimate all the sizes Δ_1, Δ_2, ...

Let us estimate these sizes one by one.

Estimating Δ_1. The only information that we have about the value Δ_1 is that it is larger than 0 and smaller than Δ. In other words, in principle, Δ_1 can take any value from the interval $(0, \Delta)$. We have no information about the probabilities of different values from this interval.

Since we have no reason to think that some specific value from this interval is more probable and some other specific value is less probable, it is reasonable to assume that all the values from this interval are equally probable. This argument – known as *Laplace indeterminacy principle* – is widely used in statistics applications; see, e.g., [1, 4]

In precise terms, this means that we assume that the distribution of possible values Δ_1 on the interval $(0, \Delta)$ is *uniform*. For the uniform distribution on an interval, the expected value is the interval's midpoint. Thus, the expected value of Δ_1 is equal to $\Delta/2$.

This expected value is what we will use as an estimate for Δ_1.

Estimating Δ_2, etc. Once we produced an estimate $\Delta_1 = \Delta_2$, a next step is to estimate the next inaccuracy component Δ_2. The only information that we have about the value Δ_2 is that it is larger than 0 and smaller than Δ_1. In other words, in principle, Δ_2 can take any value from the interval $(0, \Delta_1)$. We have no information about the probability of different values from this interval. Thus, similarly to the previous section, it is reasonable to assume that Δ_2 is uniformly distributed on the interval $(0, \Delta_1)$. In this case, the expected value of Δ_2 is equal to $\Delta_1/2$. This expected value $\Delta_2 = \Delta_1/2$ is what we will use as an estimate for Δ_2.

Similarly, we conclude that a reasonable estimate for $\Delta_3 < \Delta_2$ is $\Delta_3 = \Delta_2/2$, and, in general, that for every k, a reasonable estimate for $\Delta_{k+1} < \Delta_k$ is $\Delta_{k+1} = \Delta_k/2$.

Adding up all these estimates lead to the desired explanation. From $\Delta_1 = \Delta/2$ and $\Delta_{k+1} = \Delta_k/2$, we can conclude, by induction, that $\Delta_k = 2^{-k} \cdot \Delta$. Substituting these estimates into the formula for the overall inaccuracy $\delta = \Delta + \Delta_1 + \Delta_2 + \ldots$, we conclude that

$$\delta = \Delta + 2^{-1} \cdot \Delta + 2^{-2} \cdot \Delta + \ldots = \left(1 + 2^{-1} + 2^{-2} + \ldots\right) \cdot \Delta.$$

The sum of the geometric progression $1 + 2^{-1} + 2^{-2} + \ldots$ is 2, so we get $\delta = 2\Delta$.

This is exactly the formula that we tried to explain – with a safety factor of two.

Acknowledgments. This work was supported in part by the National Science Foundation grants HRD-0734825 and HRD-1242122 (Cyber-ShARE Center of Excellence) and DUE-0926721.

References

[1] E. T. Jaynes, *Probability Theory: The Logic of Science*, Cambridge University Press, Cambridge, Massachisetts, 2003.

[2] N. B. Maslova, Private communication.

[3] W. J. M. Rankine, *A Manual on Applied Mechanics*, Richard Griffin and Company, Glasgow, Scotland, 1858, pp. 274, available at https://archive.org/stream/manualappmecha00rankrich#page/n8/mode/1up

[4] D. J. Sheskin, *Handbook of Parametric and Nonparametric Statistical Procedures*, Chapman and Hall/CRC Press, Boca Raton, Florida, 2011.

[5] R. A. Smith and S. Hillmansen, "A brief historical overview of the fatigue of railway axles", *Proceeding of the Institute of Mechanical Engineering*, 2004, Vol. 218, pp. 267–277.

[6] D. Ullman, *Mechanical Design Process*, McGraw-Hill, New York, 2009, pp. 314–317.