Guard Digits versus Roundoff Error versus Overall Uncertainty

John Denker, Tucson, AZ
Larry Smith, Snow College, Ephraim UT

Phys. Teach. **56** 532 (November 2018) © AAPT
http://doi.org/10.1119/1.5064563

Roundoff error is an error. It can be dramatically reduced by the use of additional low-order digits, i.e. *guard digits*.[1–6] Although the significant-figures[7] idea in its standard form is incompatible with guard digits, this problem can be neatly solved by underlining the last "significant" digit, and then appending guard digits to the right of there. This clarifies the distinction between roundoff error[8] and overall uncertainty[9] while preserving compatibility with sig figs. The focus here is on reducing roundoff error; we defer discussion of most other contributions to the uncertainty.[11, 12]

The number π can be approximated in decimal notation in various ways, such as 3.14 or 3.1416. Each one differs from the true value of π by some amount we call the *quantization error* or equivalently the *roundoff error*. Similarly, when a longer decimal is rounded to make a shorter decimal, the difference is again called roundoff error.[13]

In the dictionary sense of the word, there is no "uncertainty" when rounding π, because the roundoff error is not random, and it's not even unknown. However, in physics we use a generalized notion of uncertainty, which includes anything that makes the results less than exact. This includes, among other things, roundoff error.

When using the tolerance-based approach,[14, 15] you first figure out how accurate the result needs to be to serve the intended purpose, and then design the measurements and calculations so that the result will be within tolerances. It is usually easy to keep roundoff error very small on this scale.

For example: Suppose we want to know the amount of gas in a bottle, accurate to a few percent. One student measures $P \approx 97.6$ kPa, another measures $V \approx 1.216$ L, and a third measures $T \approx 294$ K, all accurate to 1% or better. A fourth student collects those readings, looks up $R \approx 8.314\,4598$ J/K/mol, and calculates $n{=}PV/RT \approx 0.04855$ mol, which is a correct and complete answer. The students don't need to worry about uncertainty in any detail. They keep plenty of digits, to ensure that roundoff error is tiny compared to the tolerances.

Optionally, any student who is interested can use simple proportional reasoning (or Crank Three Times[11]) to verify that the experimental uncertainty inherited by the output from the raw data is well within tolerances. Also, the teacher should verify this at the time the activity is designed.

Conceptual point: So far, the roundoff error has had almost no connection to the uncertainty inherited from the raw data. They are both within tolerances, but that's about it.

Most unfortunately, sig figs blurs the distinction between roundoff and other contributions to the uncertainty. It not only assumes roundoff error is the dominant contribution, it requires you to make it so, by means of very coarse rounding.

We now turn to the metrology-style approach: In this case, you don't know the purpose of the result, so you measure things as carefully as possible, report the resulting uncertainty, and let people make of it what they will. Whereas the tolerance approach compares roundoff error to the tolerances, this approach compares it to other sources of uncertainty. Usually the roundoff error is very small on this scale. If not, keep more digits. Each additional digit reduces the maximum possible roundoff error by a factor of 10.

Guard digits is a somewhat loose term, referring to any digits that serve primarily to make roundoff error small compared to whatever you really care about: the overall uncertainty and/or the tolerance, as the case may be.

Since sig figs is featured in some introductory textbooks, let's revisit the previous calculation, using sig figs ideas to obtain a rough estimate of the uncertainty. We underline[16] the last digit that sig figs would allow, to indicate that the overall uncertainty is comparable[17] to half a count in the underlined decimal place. The digits to the right of there are guard digits. So, $P{=}9\underline{7}.6$ kPa, $V{=}1.2\underline{1}6$ L, and $T{=}2\underline{9}4$ K.[18] That gives $n{=}0.04\underline{8}55$ mol, which is represented by the magenta diamond[19] with error bars in figure 1.

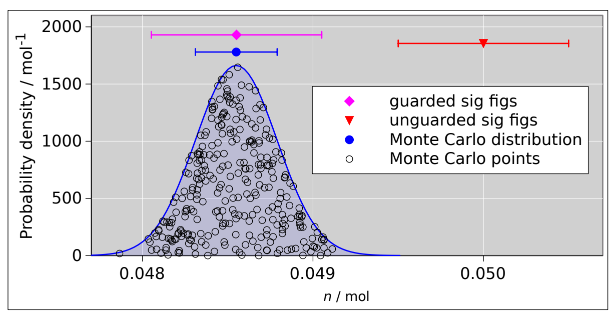

Figure 1. Various Calculations of $n = PV/RT$

If our students tried to represent uncertainty using sig figs without guard digits, they would be forced to round their data quite coarsely: $P{=}9\underline{8}$ kPa, $V{=}1.2\underline{2}$ L, and $T{=}2\underline{9}0$ K. That gives $n{=}0.05\underline{0}$ mol, which is represented by the red triangle with error bars in figure 1. It differs from the correct answer by multiple standard deviations.

The scattered black points under the bell curve were obtained via Monte Carlo, i.e. by simulating the experiment to high accuracy millions of times. The details are beyond the scope of this discussion, but the point is, there is no doubt what the right answer is. If you want to get the right answer, you need to use guard digits.

Here's another example: Suppose a cylindrical tank has a measured circumference of $3000{\pm}3$ inches, and we wish to find the radius. Dividing by 2π gives us $477.46{\pm}0.48$ inches, which is a good answer. The relative uncertainty is 0.1% on both the circumference and the radius, which makes sense.

We can redo the calculation using sig figs as follows: The circumference is $30\underline{0}0$ inches. The underline tells us there are three significant digits. (Without the underline, it would be unclear whether 3000 had one, two, three, or four significant digits. The ambiguity could have been removed by using scientific notation, but $30\underline{0}0$ is concise and convenient.) The radius is $47\underline{7}.46$ inches. Without guard digits, sig figs would require rounding to $47\underline{7}$ inches, which would, alas, displace the answer by almost an entire error bar. So don't do that. Keep at least[20, 21] one digit more than sig figs would allow.

Discussion

Guard digits are valuable for final results as well as raw observations and everything in between. In the real world you normally assume that somebody is going to *use* your results. Your final output is some colleague's input.

One consequence is that not all correct answers will be digit-by-digit identical. This reflects the fact that scatter in the data is not a mistake; it is normal and necessary. If you round the data to the point where no scatter remains, i.e. where no guard digits remain, healthy random scatter is replaced by a larger amount of unhealthy systematic bias. This is a trap for the unwary: The data looks prettier, even though it is much worse in terms of agreement with experiment. Remember: A somewhat-unreliable digit is more informative than an absent digit.

Guard digits should be integrated into every measurement and every calculation, in class and out, from now to the end of time.

When in doubt, redo the calculation with more digits. If the new answer is significantly different, it is surely better. This check can be performed once and for all at the time the procedure is designed, but students should be asked to re-check it on occasion. In particular, to drive home the importance of guard digits, you can pick a suitable[22] multi-step calculation and have students do it twice: once with too few digits at every step, and once with more.

When using a calculator, it is good practice to *leave intermediate results in the machine*. Use the memory features such as Ans, Sto, and Rcl, and/or use the RPN stack. If you wish to write something down, that's fine, but leave it in the machine as well, rather than keying it in again. This is easier and less vulnerable to typos. It keeps 15 digits or more.

When reading a digital instrument, write down all the digits. This is easier than trying to round on the fly, and preserves more information. If you know how accurate the instrument is, you should record that your lab book, but that's separate from writing down today's measurements. If possible, choose a sufficiently sensitive range, so that at least[12, 21] one low-order digit is unreliable, due to noise, nonlinearity, hysteresis, or whatever. Guard digits are valuable, and they are not expected to be reproducible.

When reading an analog instrument, start by writing down all the digits that the graduations directly provide. Then, depending on the instrument — and on your tolerances — it might or might not be worth the trouble to interpolate between graduations to estimate another digit. If you are fond of sig figs, underline the last digit that sig figs would permit, but write down *at least* one more than that.

Underlining helps sig-figs users disambiguate two key concepts: The roundoff error at each step[20] is comparable[17] to a half count in the very last digit, while the overall uncertainty is comparable to a half count in the underlined digit. Underlining is not required, except as a pedagogical stepping stone on the path that leads from sig figs toward simpler and/or more reliable methods.[11, 15] (Sig figs with guard digits but without underlining would be a recipe for endless confusion.)

There is never a penalty for "extra" accuracy. If a student writes down some huge number of guard digits, that's not a problem, and even if it were, it would soon solve itself, since there is already more than enough natural incentive for rounding off.

There is more than one right answer to the question of how many digits to keep. One guard digit is usually the minimum. Additional digits provide additional safety margin[20, 21] at very little cost, whereas keeping too few digits is often catastrophic. Especially when using a calculator or computer, keeping extra digits is usually easier than deciding whether they are necessary or not. So the take-home message is simply this:

Keep sufficiently many digits to avoid unintended loss of information. Keep sufficiently few to be reasonably convenient.

Notes and References

1. The guard-digits concept and terminology are standard in the numerical methods community, and have been for decades.[2–4] Guard digits are recommended (or at least tolerated) by some introductory physics textbooks, but forbidden by others.
2. Forman S. Acton, *Numerical Methods that Work* (Mathematical Association of America; 1st ed. 1970; updated 1990).
3. David Goldberg, "What Every Computer Scientist Should Know about Floating-Point Arithmetic" (ACM Computing Surveys, **23** 1, March 1991). `http://www.akira.ruc.dk/~keld/teaching/CAN_e14/Readings/Goldberg91.pdf`
4. Wikipedia article: "Guard digits" `https://wikipedia.org/wiki/Guard_digits`
5. R.H. Good, "Wrong rounding rule" *Phys. Teach.* **34**, 192 (1996) `http://dx.doi.org/10.1119/1.2344400`
6. Abraham Vilchis, "Significant Figures in Measurements with Uncertainty: A Working Criterion" *Phys. Teach.* **55**, 173 (2017) `http://dx.doi.org/10.1119/1.4976663`
7. Wikipedia article: "Significant figures" `https://wikipedia.org/wiki/Significant_figures`
8. Eric W. Weisstein, "Roundoff Error" (MathWorld) `http://mathworld.wolfram.com/RoundoffError.html`
9. Uncertainty is best defined in terms of probability.[10] For example, it can be quantified in terms of the standard deviation of the distribution in figure 1.
10. Andy Buffler, Saalih Allie, and Fred Lubben, "Teaching Measurement and Uncertainty the GUM Way" *Phys. Teach.* **46**, 539 (2008) `http://dx.doi.org/10.1119/1.3023655`
11. "Crank Three Times" [submitted to TPT].
12. In the research lab, in the kitchen, and elsewhere, on measuring devices where the range and sensitivity are fixed, there is usually (albeit not always) at least one built-in guard digit. More uncertainty comes from drift, hysteresis, nonlinearity, etc. than from roundoff or readability. Manufacturers could redesign the devices to provide a coarser display — but you wouldn't want them to. That would increase the overall uncertainty, making the devices less useful. Furthermore, the thing being measured is often less than perfectly reproducible. If you repeatedly measure "the" thickness of a shag carpet or a potato using a micrometer caliper, you will get many different answers. Do not discard uncertain digits just because they are uncertain! They are your guard digits.
13. When using a calculator, you shouldn't key in π as a decimal, but instead rely on the calculator's built-in notion of π, for which the roundoff error is probably 10^{-15} or less.
14. Clifford E. Swartz, "EDITORIAL: Insignificant figures" *Phys. Teach.* **6**, 125 (March 1968). `http://dx.doi.org/10.1119/1.2352406`
15. "Tolerance, Sensitivity Analysis, and Uncertainty" [submitted to TPT].
16. Wikipedia[7] suggests overlining, but that conflicts with the notation for repeating decimals.
17. Anything between 0.15 and 1.5 is considered "comparable" to a half. However, it's not worth being super-precise about it, because the sig figs representation is a very coarse approximation. It rounds the uncertainty to the nearest order of magnitude, or worse, which is not good enough for many real-world applications.[23] (The $A \pm B$ representation is better.) If/when roundoff is dominant, the worst-case error is exactly half a count, and the RMS error, averaged over all possible roundoff errors, is approximately 0.289 counts.
18. We bow to convention and write $T = 2\underline{9}4$ K, even though it would be more logical to write $T \in 2\underline{9}4$ K, using the set-membership symbol, since T is a single point whereas $2\underline{9}4$ K is a set of points, i.e. a probability distribution.[9]
19. In the figure, we focus attention on the position of the points along the n axis. The height above the axis has no immediate physical significance.
20. In multi-step calculations, including iterative calculations, you need to worry about *accumulated* roundoff error, which can be much larger than the roundoff error at any one step. This increases the number of guard digits required for intermediate steps. Also, correlations and cancellations can greatly increase the need for guard digits. Intermediate results are often correlated even when the raw data is not.
21. If the data is to be subjected to signal averaging, curve fitting, or more advanced data analysis, multiple guard digits may be needed.
22. Pick a calculation where the roundoff errors accumulate, as in the PV/RT example, not one where they fortuitously cancel.
23. We care about precise values for the uncertainty whenever one data point is to be *weighed* against others. This is important for curve fitting, decision theory, et cetera. The details are beyond the scope of this article.