Calculating Confidence Intervals and Limits

Andrew J. Pounds, Ph.D.
Department of Chemistry, Mercer University

May 25, 2009

1   Introduction

When one considers the normal distribution curve, one can claim that 68.3% of all the values fall between $\bar{x} \pm \sigma$ where σ is the *population standard deviation* based on a large number (theoretically infinite) number of trials. This interval is known as a confidence interval and its limits as confidence limits.

When only a limited number of measurements of a quantity are available, the magnitude of the estimate of the *sample standard deviation*, s, depends on the number of measurements, and consequently the calculated confidence interval also depend on that number. For N measurements, the confidence interval of a single measurement can be expressed as $\bar{x} \pm ts$. The value of the parameter t depends on the number of measurements. These values can be read off of a table or explicitly calculated.[1]

2   Example

Assume that one has the following set of data shown in Table 1. Then $\bar{x}$ is calculated as...

Table 1: Sample Data

i	x
1	20.02
2	20.11
3	21.05
4	20.66
5	19.59
6	20.87

$$\bar{x} = \frac{\sum_{i=1}^{N} x_i}{N} = \frac{20.02 + 20.11 + 21.05 + 20.66 + 19.59 + 20.87}{6} = 20.38 \qquad (1)$$

The sample standard deviation is given by (Eqn. 2).

$$s = \sqrt{\frac{\sum_{i=1}^{N}(x_i - \bar{x})^2}{N-1}} \qquad (2)$$

---

[1] Flashka, Barnard, and Sturrock. *Quantitative Analytical Chemistry, 2nd ed.* (Boston: Willard Grant Press, 1980), p.14

Using Eqn 2, the data from Table 1, and the results from Eqn. 1, one computes the sample standard deviation as follows.

$$
\begin{aligned}
s &= \sqrt{\frac{\sum_{i=1}^{N}(x_i - \bar{x})^2}{N-1}} \\
&= \sqrt{\frac{(20.02 - 20.38)^2 + \cdots + (20.86 - 20.38)^2}{6-1}} \\
&= \sqrt{\frac{1.594}{5}} \\
&= 0.565
\end{aligned}
$$

The 95% confidence interval would then be calculated by first looking up (or calculating) the t value for a confidence interval of 95% and 5 degrees of freedom (that's N-1). The correct value, to 4 significant figures, is 2.015. The correct way to list the measured value at a 95% confidence interval is to write it with its proper confidence limits using the number of significant digits in the measured quantity as to determine the precision of the confidence interval. For example

$$
\begin{aligned}
\bar{x} &\pm ts \\
20.38 &\pm (2.015)(0.565) \\
20.38 &\pm 1.14
\end{aligned}
$$

The 20.38 ± 1.14 is the experimental value with its 95% confidence limit explicitly represented.

Another less commonly used practice is to let the confidence interval actually determine the digits of precision represented in the final result. While this technique is more widely accepted using standard deviations, it can be done with confidence intervals as well. When large sets of data are collected, one can round the standard deviation to one significant figure and use the magnitude of the standard deviation to determine the number of significant figures for the final result.[2] By analogy, if one has a smaller sample size (where t-test statistics are more applicable) one can use the confidence interval to determine the number of digits of precision to represent in the final result. In the example below, the confidence interval is rounded to one digit and that is used as a means to both round the average and also determine the number of significant digits to which it can be represented.

$$
\begin{aligned}
\bar{x} &\pm ts \\
20.38 &\pm (2.015)(0.565) \\
20.38 &\pm 1.14 \\
20 &\pm 1
\end{aligned}
$$

So in this example, the final result would only have two significant digits.

3 Calculating t Values

While you will not be expected to explicitly calculate t value,[3] some might find the method by which they are calculated interesting. The t value for a given confidence interval, C, and the numbers of degrees of

---

[2]This does not necessarily apply to the case when the quantity is a derived quantity. In those cases methods of error propagation have to be applied.

[3]For most standard confidence intervals these are tabulated. You can also use the online tool at http://theochem.mercer.edu/chm112/ttable.

freedom, ν, must satisfy the following equation.

$$C = \int_{-\infty}^{t} \frac{\Gamma\left((\nu+1)/2\right)}{\sqrt{\nu\pi}\Gamma\left(\nu/2\right)\left(1+x^2/\nu\right)^{(\nu+1)/2}} dx \tag{3}$$

Where Γ refers to the Gamma Function. To solve this equation for t, the equation was rewritten as

$$g(t, C, \nu) = C - \int_{-\infty}^{t} \frac{\Gamma\left((\nu+1)/2\right)}{\sqrt{\nu\pi}\Gamma\left(\nu/2\right)\left(1+x^2/\nu\right)^{(\nu+1)/2}} dx \tag{4}$$

and the parameters C and ν were fixed. After setting $g(t, C, \nu) = 0$, numerical methods of quadrature and root finding were used to find the value of t to satisfy the equation.