Biological Chemistry Laboratory
Biology 3515/Chemistry 3515
Spring 2022

Lecture 6:
Measuring Protein Concentration with UV-visible Spectrophotometry
and
Dealing with Uncertainty

Thursday, 27 January 2022

©David P. Goldenberg
University of Utah
goldenberg@biology.utah.edu

Computer Labs

- Computer Labs next week and the following week.
  - Start at 2:00 PM!
  - Room 150 S. Biology Building
- Next week: Graphing and curve fitting with SciDAVis.
- Following week: Molecular modeling with PyMOL.
- We will use the computers in the lab, not personal laptops.
- But, you should still install SciDAVis and PyMOL on your own computer. Use the versions available on Canvas.

Direct Methods for Measuring Protein Concentration by Absorbance

1. Direct measurement of UV absorbance (usually at 280 nm)
   - Very useful for pure protein samples, but need to know the extinction coefficient.
   - Extinction coefficient is specific to the protein and depends primarily on the number of Tyr and Trp residues per molecule.

     Can be estimated reasonably well from the amino acid sequence or composition.
   - Not especially sensitive. Good for concentrations of $\approx$ 0.1 mg/mL or greater.
   - Absorbance from other compounds can interfere.
2. Direct measurement of visible absorbance.
   - Very useful for metalloproteins containing Fe or Cu.
   - Need to know extinction coefficient.

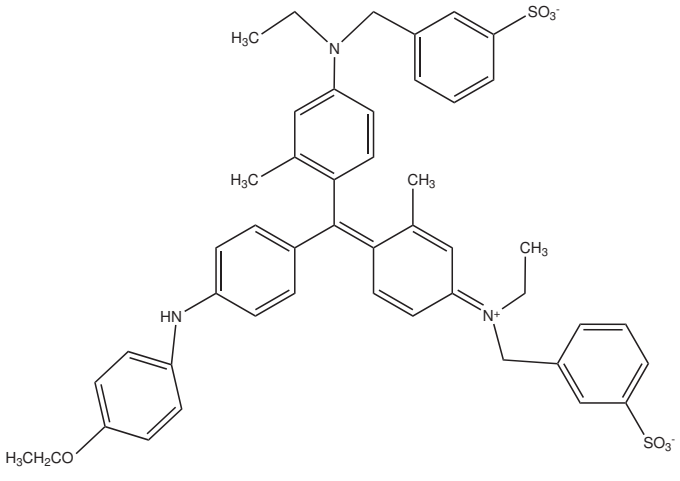
Indirect Methods for Measuring Protein Concentration

1. Formation of coordinated metal complexes, especially Cu.
2. Binding to dyes, leading to spectral shift of the dye.

Advantages

- Much more sensitive (10 $\times$ or more) than direct UV absorbance.
- Absorbance at longer (visible) wavelengths is less sensitive to interference from other compounds than UV absorbance.

The Bradford Dye-Binding Assay



Coomassie blue G-250

- Extensive conjugation leads to absorption of visible light.
- For dye: $\lambda_{max} = 465\,\text{nm}$
- For dye bound to protein: $\lambda_{max} = 595\,\text{nm}$
- Absorption at 595 nm increases as protein is added to fixed amount of dye.
- Requires calibration for a particular batch of dye and solution conditions.

Outline of Experiment

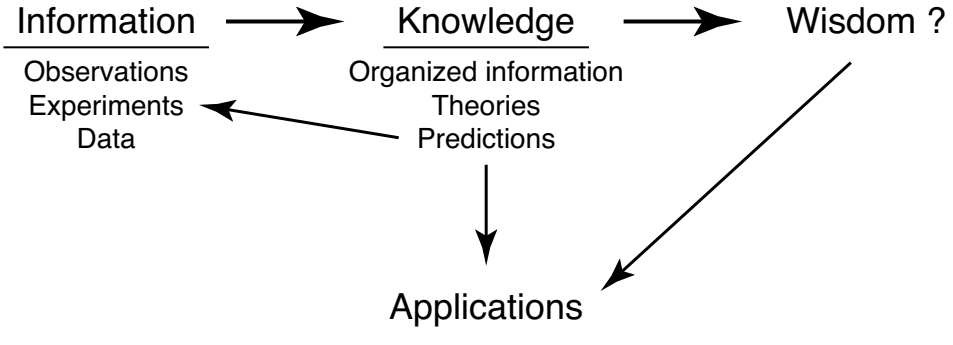
- Two samples:
  - A pure protein: Bovine serum albumin (BSA)
  - An *E. coli* extract, containing lots of proteins and nucleic acids
- Direct UV absorbance measurements at 260 and 280 nm
  - For BSA, estimate [Protein] from A_{280} and known extinction coefficient.
  - For both samples, estimate [Protein] and [NA] from extinction coefficients for "typical" proteins and nucleic acids.
- Bradford dye-binding assay
  - Use BSA to establish a standard curve, using [BSA] determined from A_{280}
  - Independent estimate [Protein] in *E. coli* extract, to be compared with estimate from $A_{280} : A_{260}$

Warning!

Direction Change

Dealing with Uncertainty

How do we know? What do we do with it?



- All of this is sometimes messy!

Dealing With Uncertainties

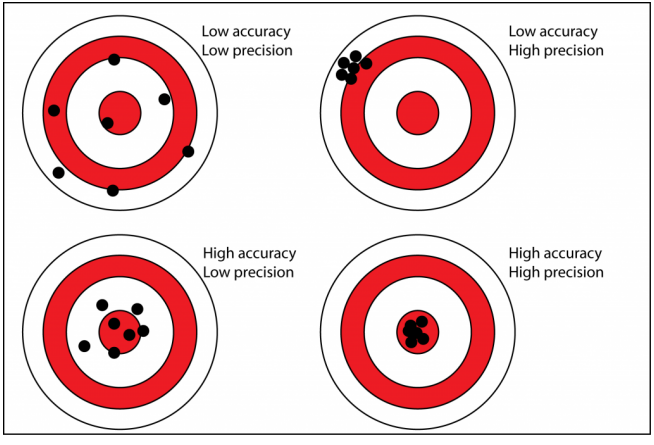
Pipette calibration data:

- Mass of water (mg) delivered from a pipette set to 20 μL:

  20.1
  18.5
  18.2
  22.4
  22.9

- mean (average) = 20.4 mg
- What is the significance of the mean?
- How do we quantify accuracy or precision?

Precision and Accuracy as Target Practice



http://www.antarcticglaciers.org/glacial-geology/dating-glacial-sediments-2/precision-and-accuracy-glacial-geology/

Precision and Accuracy in Measurement

Precision

- Reproducibility of individual measurements.
- Determined by making multiple measurements and comparing them.

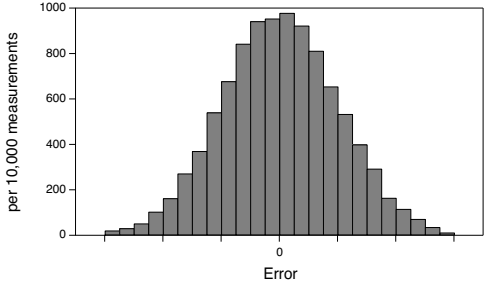
Accuracy

- Consistency with an accepted value.
- Requires comparison with an accepted standard.

Dealing With Uncertainties: The Working Model

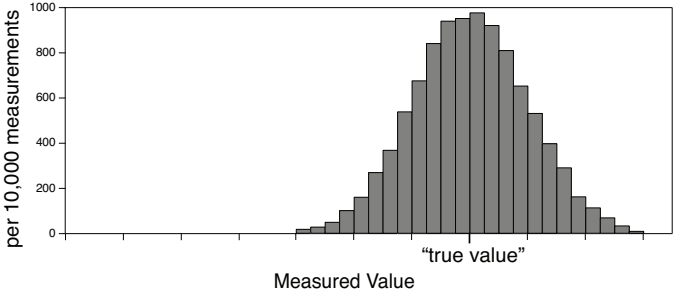
Assumptions:

- The measured values are determined by a “true” value plus random error (positive or negative).
- The random errors are distributed according to a Gaussian function, *i.e.,* a “bell curve”.



- Why is it bell shaped?

Estimating the “True” Value



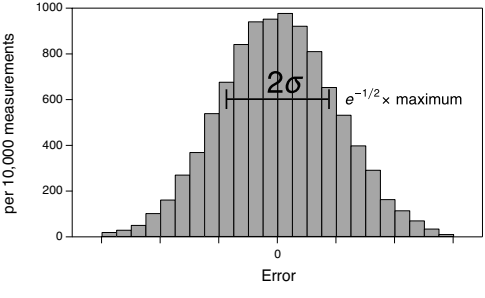
- The best* estimate of the “true” value is the mean, $\bar{x}$.

$$\bar{x} = \frac{1}{N} \sum_{i=1}^{N} x_i$$

N= number of measurements, x_i is the i^{th} measurement

* “Best” means most likely to give the correct value.

Estimating the Distribution Width (σ)

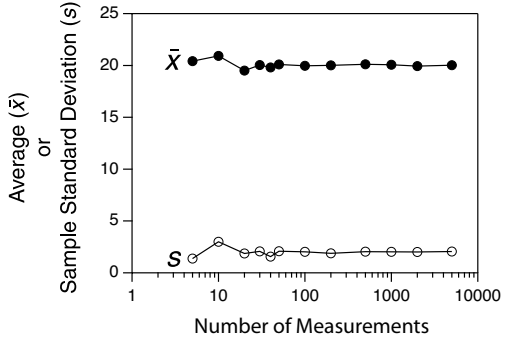


Two ways to estimate σ

- From a histogram (takes lots of measurements!)
- The sample standard deviation, s:

$$s = \sqrt{\frac{\sum (x - \bar{x})^2}{N - 1}}$$ an estimate of σ

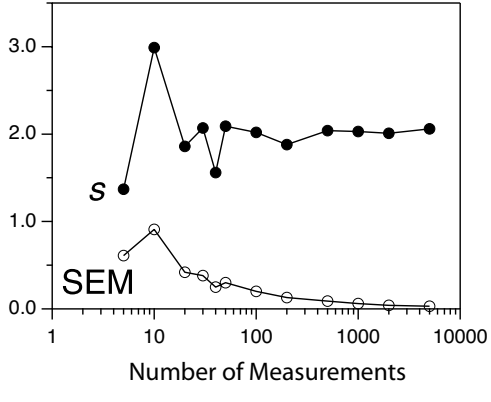
Estimates Improve With More Measurements (A Simulation)



- Estimate of true value ($\bar{x}$) approaches a limiting value (20 mg)
- Estimate of standard deviation (s) approaches a limiting value (2 mg)
- s doesn't approach zero.

Another Useful Statistic: The Standard Error of the Mean (SEM)

$$\text{SEM} = \sqrt{\frac{\sum(x-\bar{x})^2}{(N-1)N}} = s/\sqrt{N}$$



- The standard error of the mean represents the uncertainty in the estimate of the mean, $\bar{x}$
- The uncertainty in $\bar{x}$ decreases with more measurements.
- The uncertainty in the mean can be made as small as we like, if we make enough measurements! (Assumes that errors are truly random.)
- Decreasing the uncertainty by half requires four times as many measurements.

Clicker Question #1

If I want to report on how reproducible my pipette (and technique) is, which statistic should I use?

A) The sample standard deviation

B) The standard error of the mean

Clicker Question #2

If I want to report on how reliably I have measured the average volume delivered by my pipette, which statistic should I use?

A) The sample standard deviation

B) The standard error of the mean

- Whatever you report, be clear! (and specify *N*)

Significant Figures

- The basic idea: The number of digits used to report a measurement should reflect the precision of the measurement.
- Reporting more digits than justified by the measurements is dishonest!
- A precise definition of ‘significant figures’ is not so simple!

---

https://en.wikipedia.org/wiki/Significant_figures

Rules for Significant Figures

- All non-zero digits are significant.

number	sig. figs.
12	2
12.5	3

- Zeros between non-zero digits are significant.

number	sig. figs.
102	3
12.05	4

Rules for Significant Figures

- Trailing zeros to the right of a decimal point are significant.

number	sig. figs.
12.00	4
12.500	5

- Leading zeros to the left are *not* significant.

number	sig. figs.
012	2
0.0012	2

- What about trailing zeros without a decimal point?

number	sig. figs.
1200	2?

Rules for Significant Figures

- Avoid Ambiguity with Scientific Notation

number	sig. figs.
1200	2?
1.2×10^3	2
1.20×10^3	3
1.200×10^3	4
1200.	4

Rules for Significant Figures

- Numbers with unlimited significant figures:
  - Integers or ratios of integers (rational numbers), such as 2, $1/2$ or $2/3$.
  - Defined irrational numbers, such as $\sqrt{2}$, π or e.
  - Other numbers that are not derived from measurements, including most conversion factors.

Rules for Significant Figures

- Multiplication and division:

  The calculated result should contain the number of significant figures of the measured quantity with the smallest number of significant figures.

$$15\,\mathrm{g} \div 121.1\,\mathrm{g/mol} = 0.12\,\mathrm{mol}$$

$$\begin{aligned} 15\,\mathrm{mM} \times 25\,\mu\mathrm{L} &= 0.015\,\mathrm{moles/L} \times 2.5 \times 10^{-5}\,\mathrm{L} \\ &= 3.8 \times 10^{-7}\,\mathrm{moles} \\ &= 0.38\,\mu\mathrm{moles} \end{aligned}$$

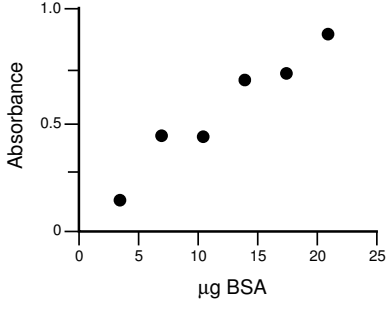
Rules for Significant Figures

- For addition and subtraction:
  - The last decimal place of the result is determined by last decimal place of the measured quantity with the smallest number of decimal places.

    $$125\,\mathrm{g} + 0.035\,\mathrm{g} = 125\,\mathrm{g}$$

  - Adding a more precise value to a less precise one doesn't increase the precision of the sum!

The Curve-Fitting Problem



- How do we find the equation of the line (or other function) that best “fits” the experimental data?
- What assumptions do we make when fitting data to a function?
- How we determine how well the function (model) fits the data?