Must Complete your Radiation Certification before next lab.

Feedback from the first labs

Linear fits

Fitting non-linear functions

Correlation and covariance
• The first labs will be returned this week (except for late submissions).

• Grades averaged about 15/20.

• Please read the comments! If you have any questions, see your section instructor.

• You have 3 weeks before next due date, but if you aren’t doing a 3-week lab now you should finish the report early to avoid conflicts with your next lab.

• Sign up next week for your third lab.
COMMON PROBLEMS IN LAB REPORTS

- Units and uncertainties: Nearly all measurements should have units and uncertainties quoted. If you believe that the uncertainty is negligible, note that.
- Technique section: Discuss the relevant physics and explain the equations.
- Apparatus: Describe, don’t just draw, the equipment.
- Logbook: Need more than just a table of numbers. Include configuration description, what experiment you are doing, and basic narrative of data collection.
- Analysis: You must describe your data analysis in the text of your report. “Data analysis: see attached computer output” is not acceptable. Include text, equations that describe the analysis steps (not just numbers and Mathcad output).
COMMON PROBLEMS IN LAB REPORTS

- Error propagation: Often inadequate, sometimes missing.
- Conclusion: Several students didn’t write scientific conclusions! Must summarize results, including important numerical values. Example: “We measured e/m for the electron to be xx ± yy. Dominant uncertainties were from....” Show your result with uncertainty, compare to expected value, and quote discrepancy in sigma, not as percent discrepancy. The latter is meaningless without an understanding of the measurement error. Note that error is not discrepancy. Discuss whether there were likely other errors that you didn’t account for.
- Page ordering: Please start your report with the introduction. Raw data sheets are supplemental and should be in an appendix.
STANDARD DEVIATION AND MEASUREMENT UNCERTAINTY

• Measure something N times; take the average of the measurements to find best estimate $X$ of the true value.

• Calculate standard deviation $\sigma$

• What is the statistical error on your result?
  • $\sigma$?
  • $\sigma/\sqrt{N}$?
  • Something else?

• Your measurement is the mean — so your error has to be the error on the mean.
USING STANDARD DEVIATION: EXAMPLE (FRANCK-HERTZ)

- Meter has systematic error of 0.01 V
- What is the uncertainty?
  - Calculate standard deviation of the 10 measurements: \( \sigma = 0.134 \) V
  - Statistical error on mean is \( \frac{\sigma}{\sqrt{10}} = 0.042 \) V
  - This is way bigger than systematic error from meter error (of course there may be other systematic errors!).
- So we can write that \( V_{\text{ion}} = (11.46 \pm 0.04) \) V
TURNING MAXIMUM LIKELIHOOD ARGUMENT INTO A WEIGHTED AVERAGE

- Take N measurements $x_1, x_2, \ldots, x_N$ with uncertainties $\sigma_1, \sigma_2, \ldots, \sigma_N$
- Most likely true value given by weighted average
- The weights $w_i$ are $w_i = \frac{1}{\sigma_i^2}$
- The uncertainty of $x_{ave}$ is $\sigma_{ave} = \frac{1}{\sqrt{\sum w_i}}$

- If $\sigma_1 = \sigma_2 = \ldots = \sigma_N$, weighted average = straight average

Weights go as $1/\sigma^2$, so more precise measurements count much more than less precise ones.
WEIGHTED AVERAGE: WHEN?

- **NO**, if:
  - You don’t know the intrinsic uncertainty, or
  - The intrinsic uncertainty is the same for all measurements
  - You need to calculate the standard deviation to determine the uncertainty

- **YES**, if:
  - You know the intrinsic uncertainties, AND
  - They are not the same for all measurements
LINE FITS

- Weighted average used to estimate single parameter (assume all points measuring same quantity)
- Sometimes have data points which should lie on line
- Example is photoelectric effect
- $h \nu = T_{\text{max}} + \phi$ and $eV_s = T_{\text{max}}$
- $V_s = \frac{h}{e} \nu - \frac{\phi}{e}$
- Plot $V_s$ versus $\nu$
- Fit to a line to find slope $(h/e)$ and work function $\phi$
- Actual experiment more complicated: $\frac{\phi}{e} \to \frac{\phi}{e} - V_c$
LINE FITS

- Measure $N$ points with values $(x_1, y_1), (x_2, y_2), \ldots, (x_N, y_N)$ with negligible uncertainty on $x$ and constant $\sigma_y$ uncertainty on all $y$ values.

- Assume measurements lie on line with parameters $y = A + Bx$.

- For a given $x_i$, true value of $y_i$ is $A + Bx_i$ (instead of $\mu$).

- Probability of obtaining $y_i$ is $\propto \frac{1}{\sigma_y} \exp \left[ -\frac{(y_i - (A + Bx_i))^2}{2\sigma_y} \right]$.

- Probability of obtaining complete set of measurements is $\propto \frac{1}{\sigma_y^N} \exp \left[ -\frac{\chi^2}{2} \right]$ where $\chi^2 = \sum \frac{(y_i - A - Bx_i)^2}{\sigma_y^2}$.

- Maximum probability when $\chi^2$ is a minimum.

- Find minimum by differentiating $\chi^2$ with respect to $A$ and $B$ and set equal to 0.
\[ \frac{\partial \chi^2}{\partial A} = -\frac{2}{\sigma_y^2} \sum_{i=1}^{N} (y_i - A - Bx_i) = 0 \]
\[ \frac{\partial \chi^2}{\partial B} = -\frac{2}{\sigma_y^2} \sum_{i=1}^{N} x_i(y_i - A - Bx_i) = 0 \]

can be rewritten as:

\[ AN + B \sum x_i = \sum y_i \]
\[ A \sum x_i + B \sum x_i^2 = \sum x_i y_i \]

which can be solved for:

\[ A = \frac{\sum x_i^2 \sum y_i - \sum x_i \sum x_i y_i}{\Delta} \]
\[ B = \frac{N \sum x_i y_i - \sum x_i \sum y_i}{\Delta} \]

where \( \Delta = N \sum x_i^2 - (\sum x_i)^2 \)

Problem 8.9, page 201 gives results for cases when uncertainties on \( y_i \) are not all equal (weighted linear fit)
Our standard deviation from many measurements was:

\[ \sigma_x = \sqrt{\frac{\sum_{i=1}^{N} (x_i - \overline{x})^2}{N - 1}} \]

The factor of \( N - 1 \) instead of \( N \) came from using \( \overline{x} \) instead of (unknown) \( \mu \).

Similar result for estimating the uncertainty in \( y \):

\[ \sigma_y = \sqrt{\frac{\sum_{i=1}^{N} (y_i - A - Bx_i)^2}{N - 2}} \]

Here the expected value for \( y_i \) is \( A + Bx_i \) instead of \( \overline{x} \) and we have \( N - 2 \) instead of \( N \) due to guessing on \( A \) and \( B \) (for \( N = 2 \) we are guaranteed to have a perfect line and therefore no information on uncertainties).
FITTING A LINE

- Example from photoelectric effect calibration:

Measure voltage on capacitor by discharging through galvanometer converting charge to displacement so $V = \frac{1}{kC} D$ where $V$ is calibration voltage, $D$ is galvanometer deflection, and $1/kC$ converts $D$ to $V$. Here $V$ is independent ($x$) and $D$ is dependent ($y$) so use $D = kC \cdot V$

<table>
<thead>
<tr>
<th>Voltage (V)</th>
<th>Deflection (cm)</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>0.00</td>
</tr>
<tr>
<td>0.2</td>
<td>1.1</td>
</tr>
<tr>
<td>0.4</td>
<td>2.3</td>
</tr>
<tr>
<td>0.6</td>
<td>3.7</td>
</tr>
<tr>
<td>0.8</td>
<td>4.5</td>
</tr>
<tr>
<td>1.0</td>
<td>6.0</td>
</tr>
<tr>
<td>1.2</td>
<td>7.3</td>
</tr>
<tr>
<td>1.4</td>
<td>8.5</td>
</tr>
<tr>
<td>1.6</td>
<td>9.4</td>
</tr>
<tr>
<td>1.8</td>
<td>10.9</td>
</tr>
<tr>
<td>2.0</td>
<td>12.1</td>
</tr>
</tbody>
</table>
FITTING A LINE: ALL ERRORS THE SAME

Use least-squares formula from Taylor Sec. 8.2 to fit the N measurements to the form \( y = A + Bx \), where \( x \leftrightarrow V \) and \( y \leftrightarrow D \) (assume that errors on \( x \) are negligible compared to errors on \( y \)):

\[
A = \frac{\sum x_i^2 \sum y_i - \sum x_i \sum x_i y_i}{\Delta}, \quad B = \frac{N \sum x_i y_i - \sum x_i \sum y_i}{\Delta}, \quad \Delta = N \sum x_i^2 - (\sum x_i)^2
\]

\[
\sum x_i = 0.0 + 0.2 + 0.4 \ldots 2.0 = 11.0 \text{ V}
\]

\[
\sum x_i^2 = 0.0 + 0.2^2 + 0.4^2 \ldots 2.0^2 = 15.4 \text{ V}^2
\]

\[
\sum y_i = 0.0 + 1.1 + 2.3 \ldots 12.1 = 65.8 \text{ cm}
\]

\[
\sum x_i y_i = 0.0 \cdot 0.0 + 0.2 \cdot 1.1 + 0.4 \cdot 2.3 \ldots 2.0 \cdot 12.1 = 92.48 \text{ V} \cdot \text{cm}
\]

\[
\Delta = N \sum x_i^2 - (\sum x_i)^2 = 11 \times 15.4 - 11.0^2 = 48.4 \text{ V}^2
\]

\[
A = \frac{\sum x_i^2 \sum y_i - \sum x_i \sum x_i y_i}{\Delta} = \frac{15.4 \times 65.8 - 11.0 \times 92.48}{48.4} = -0.08 \text{ cm}
\]

\[
B = \frac{N \sum x_i y_i - \sum x_i \sum y_i}{\Delta} = \frac{11 \times 92.48 - 11.0 \times 65.8}{48.4} = 6.06 \text{ cm/V} = kC
\]
FITTING A LINE

From spread of measurements, find uncertainty on $y$, $A$, and $B$:

$$\sigma_y = \sqrt{\frac{\sum(y_i-A-Bx_i)^2}{N-2}} = 0.14\text{ cm} = \sigma_D$$

$$\sigma_A = \sigma_y \sqrt{\frac{\sum x_i^2}{\Delta}} = 0.14 \times \sqrt{\frac{15.4}{48.4}} = 0.08\text{ cm}$$

$$\sigma_B = \sigma_y \sqrt{\frac{N}{\Delta}} = 0.14 \times \sqrt{\frac{11}{48.4}} = 0.07\text{ cm/V}$$

A is consistent with zero: this is a good check, since in this case the intercept should be zero as $D$ should be directly proportional to $V$.

Calibration constant $kC$:

$$kC = (6.06 \pm 0.07)\text{ cm/V}$$
FITTING TO AN EXPONENTIAL

- Some functions can be easily “converted” to linear and then fitted to a line.
- Exponential decay shows up in many places, especially radioactive decays.
- Decay rate drops with time: $N(t) = N_0 \exp(-t/\tau)$ where $\tau$ is the mean lifetime; half-life is related by factor of $\ln(2)$.
- However, $\ln(N) = \ln(N_0) - t/\tau$ which is linear!
- So, can do a linear fit to $\ln(N)$ and extract $\tau$. 
FITTING TO AN EXPONENTIAL

- Start with \( N(t) = N_0 e^{-\lambda t} \)
- We want the decay constant \( \lambda \)
- Take natural logarithm to get \( \ln N = \ln N_0 - \lambda t \)
- Identify \( y \) as \( \ln N \), \( x \) as \( t \), \( A \) as \( \ln N_0 \) and \( B \) as \( \lambda \)
- Use previous formulas to find \( \lambda \)
FITTING TO A MORE GENERAL FUNCTION

• Possible to fit a completely general function $f(x)$ where $f$ is dependent on parameters $A$, $B$, $C$, ...

• Obtain $N$ measurements $y_i$ for $N$ values of $x$

• Construct the chi-squared

\[ \chi^2 = \sum \frac{(y_i - f(x_i))^2}{\sigma_i^2} \]

• Scan over possible values of the parameters $A$, $B$, $C$, ... and find the values that cause $\chi^2$ to be minimized

• These values are the most likely values of the parameters. Uncertainties on the parameters are determined from how far you have to scan away from the best values to cause $\chi^2$ to increase by 1.
How do we know if we got a good fit (i.e., does the function $f$ describe the data adequately)?

Can ask how small the minimum $\chi^2$ value was. Smaller values indicate a better fit (fewer, smaller deviations between the data points and the best fit function).

Actual numerical expectations for $\chi^2$ value depends on the number $N$ of data points and the number of parameters ($A, B, C, ...$) in the function $f$. More on this in Chapter 12.