Basic Error Analysis

Physics 401
Spring 2015
Eugene V Colla
Agenda

• Errors and uncertainties
• The Reading Error
• Accuracy and precession
• Systematic and statistical errors
• Fitting errors
• Appendix. Working with oil drop data
$T = 63^\circ F \pm ?$  

Best guess $\Delta T \sim 0.5^\circ F$

Wind speed 4mph $\pm ?$  

Best guess $\pm 0.5 mph$
What and when we need to know about errors. Industry.

Clearance fit
Measurement of the speed of the light

1675 Ole Roemer: 220,000 Km/sec

Does it make sense?
What is missing?

NIST Bolder Colorado $c = 299,792,456.2 \pm 1.1$ m/s.
We do not care about accuracy better than 1mm

If ruler is not okay, we need to use digital caliper

Probably the natural limit of accuracy can be due to length uncertainty because of temperature expansion. For 53mm $\Delta L \cong 0.012mm/K$

Reading Error $= \pm \frac{1}{2}$ (least count or minimum gradation).
Fluke 8845A multimeter

Example Vdc (reading)=0.85V
\[ \Delta V = 0.83 \times (1.8 \times 10^{-5}) + 1.0 \times (0.7 \times 10^{-5}) \approx 2.2 \times 10^{-5} \]
\[ = 22 \mu V \]

8846A Accuracy

Accuracy is given as \( \pm (\% \ measurement + \% \ of \ range) \)

<table>
<thead>
<tr>
<th>Range</th>
<th>24 Hour (23 ( \pm )1 ( ^\circ )C)</th>
<th>90 Days (23 ( \pm )5 ( ^\circ )C)</th>
<th>1 Year (23 ( \pm )5 ( ^\circ )C)</th>
<th>Temperature Coefficient/ ( ^\circ )C Outside 18 to 28 ( ^\circ )C</th>
</tr>
</thead>
<tbody>
<tr>
<td>100 mV</td>
<td>0.0025 + 0.003</td>
<td>0.0025 + 0.0035</td>
<td>0.0037 + 0.0035</td>
<td>0.0005 + 0.0005</td>
</tr>
<tr>
<td>1 V</td>
<td>0.0018 + 0.0006</td>
<td>0.0018 + 0.0007</td>
<td>0.0025 + 0.0007</td>
<td>0.0005 + 0.0001</td>
</tr>
<tr>
<td>10 V</td>
<td>0.0013 + 0.0004</td>
<td>0.0018 + 0.0005</td>
<td>0.0024 + 0.0005</td>
<td>0.0005 + 0.0001</td>
</tr>
<tr>
<td>100 V</td>
<td>0.0018 + 0.0006</td>
<td>0.0027 + 0.0006</td>
<td>0.0038 + 0.0006</td>
<td>0.0005 + 0.0001</td>
</tr>
<tr>
<td>1000 V</td>
<td>0.0018 + 0.0006</td>
<td>0.0031 + 0.001</td>
<td>0.0041 + 0.001</td>
<td>0.0005 + 0.0001</td>
</tr>
</tbody>
</table>
The accuracy of an experiment is a measure of how close the result of the experiment comes to the true value.

Precision refers to how closely individual measurements agree with each other.
Accuracy and precision

- Not Precise, Not Accurate
- Precise, Not Accurate
- Accurate, Not Precise
- Accurate, Precise
- Errors
Systematic and random errors

- **Systematic Error**: reproducible inaccuracy introduced by faulty equipment, calibration or technique.

- **Random errors**: Indefiniteness of results due to finite precision of experiment. Measure of fluctuation in result after repeatable experimentation.

Systematic errors

Sources of systematic errors: poor calibration of the equipment, changes of environmental conditions, imperfect method of observation, drift and some offset in readings etc.

Example #1: measuring of the DC voltage

U = R * I

U = \frac{R * I - \left( \frac{R}{R_{in}} \right) E_{off}}{1 + \left( \frac{R}{R_{in}} \right)}

E_{off} = f(time, temperature)
Example #3: poor calibration

Measuring of the speed of the second sound in superfluid He4

Published data

\[ T_\lambda = 2.17 \text{K} \]

P403 results

\[ T_\lambda = 2.1 \text{K} \]

LHe

Resonator

Temperature sensor

HP34401A DMM

10\mu A

DT-470/471-50

LakeShore
Random errors

Result of measurement

\[ X_{\text{meas}} = X_{\text{true}} + e_s + e_r \]

Correct value

Systematic error

Random error

\[ e_s = 0 \]

\[ P \]

\[ x_i \]

\[ x_{\text{true}} \]

\[ e_s \]

\[ P \]

\[ x_{\text{true}} \]

physics 401
Random errors. Poisson distribution

\[ P_n(t) = \frac{(rt)^n}{n!} e^{-rt} \quad n = 0, 1, 2, \ldots \]

- \( r \): decay rate [counts/s]
- \( t \): time interval [s]
- \( P_n(rt) \): Probability to have \( n \) decays in time interval \( t \)

A statistical process is described through a Poisson Distribution if:

- **random process** → for a given nucleus probability for a decay to occur is the same in each time interval.
- **universal probability** → the probability to decay in a given time interval is same for all nuclei.
- **no correlation between two instances** (the decay of one nucleus does not change the probability for a second nucleus to decay.)
Poisson distribution

\[ P_n(t) = \frac{(rt)^n}{n!} e^{-rt} \quad n = 0, 1, 2, \ldots \]

- \( r \): decay rate [counts/s]
- \( t \): time interval [s]
- \( P_n(rt) \): Probability to have \( n \) decays in time interval \( t \)

Properties of the Poisson distribution:

- \( \sum_{n=0}^{\infty} P_n(rt) = 1 \), probabilities sum to 1
- \( \langle n \rangle = \sum_{n=0}^{\infty} n \cdot P_n(rt) = rt \), the mean
- \( \sigma = \sqrt{\sum_{n=0}^{\infty} (n-\langle n \rangle)^2 P_n(rt)} = \sqrt{rt} \), standard deviation
Poisson distribution at large $rt$

$$P_n(t) = \frac{(rt)^n}{n!} e^{-rt} \quad n = 0, 1, 2, \ldots$$

Poisson and Gaussian distributions

Carl Friedrich Gauss (1777–1855)

Gaussian distribution: continuous

$$P_n(x) = \frac{1}{\sigma \sqrt{2\pi}} e^{-(x-\bar{x})^2/2\sigma^2}$$
Error in the mean is given as \( \frac{\sigma}{\sqrt{N}} \)

\[
P_n(x) = \frac{1}{\sigma \sqrt{2\pi}} e^{-\frac{(x-\bar{x})^2}{2\sigma^2}}
\]
Measurement in presence of noise

Source of noisy signal

Actual measured values

4.89855
5.25111
2.93382
4.31753
4.67903
3.52626
4.12001
2.93411

Expected value 5V
Measurement in presence of noise

**Graphs:**
- **Top Left:** Histogram with a peak at $x = 5.2924\text{V}$, labeled $10^4$.
- **Top Right:** Histogram with a peak at $x = 5.139\text{V}$, labeled $100$.
- **Bottom Left:** Histogram with a peak at $x = 4.992\text{V}$, labeled $10^4$.
- **Bottom Right:** Histogram with a peak at $x = 5.003\text{V}$, labeled $10^6$.
For $N=10^6$ $U=4.999\pm0.001$ 0.02% accuracy
**Ag β decay**

\[ ^{108}\text{Ag} \quad t_{1/2} = 157\text{s} \]

\[ ^{110}\text{Ag} \quad t_{1/2} = 24.6\text{s} \]

Model: ExpDec2

Equation:

\[
y = A_1 \exp\left(-\frac{t}{t_1}\right) + A_2 \exp\left(-\frac{t}{t_2}\right) + y_0
\]

Reduced Chi-Sqr: 1.43698

Adj. R-Square: 0.96716

<table>
<thead>
<tr>
<th>C</th>
<th>Value</th>
<th>Standard Error</th>
</tr>
</thead>
<tbody>
<tr>
<td>C</td>
<td>(y_0)</td>
<td>0.02351</td>
</tr>
<tr>
<td>C</td>
<td>(A_1)</td>
<td>104.87306</td>
</tr>
<tr>
<td>C</td>
<td>(t_1)</td>
<td>177.75903</td>
</tr>
<tr>
<td>C</td>
<td>(A_2)</td>
<td>710.01478</td>
</tr>
<tr>
<td>C</td>
<td>(t_2)</td>
<td>30.32479</td>
</tr>
</tbody>
</table>

**Gauss fit**

Equation:

\[
y = y_0 + \frac{A}{w \sqrt{\pi/2}} \exp\left(-\frac{2((x-xc)/w)^2}{1}\right)
\]

Reduced Chi-Sqr: 4.77021

Adj. R-Square: 0.93464

<table>
<thead>
<tr>
<th>Counts</th>
<th>Value</th>
<th>Standard Error</th>
</tr>
</thead>
<tbody>
<tr>
<td>Counts</td>
<td>(y_0)</td>
<td>1.44204</td>
</tr>
<tr>
<td>Counts</td>
<td>xc</td>
<td>1.49992</td>
</tr>
<tr>
<td>Counts</td>
<td>w</td>
<td>5.93398</td>
</tr>
<tr>
<td>Counts</td>
<td>A</td>
<td>219.24559</td>
</tr>
<tr>
<td>Counts</td>
<td>(\sigma)</td>
<td>2.96699</td>
</tr>
<tr>
<td>Counts</td>
<td>FWHM</td>
<td>6.98673</td>
</tr>
<tr>
<td>Counts</td>
<td>Height</td>
<td>29.4798</td>
</tr>
</tbody>
</table>
Fitting. Analysis of the residuals

Ag β decay

Test 1. Fourier analysis

No pronounced frequencies found
Fitting. Analysis of the residuals

Ag $\beta$ decay

Test 1. Autocorrelation function

Correlation function

$y(m) = \sum_{n=0}^{\infty} f(n) g(n-m)$

autocorrelation function

$y(m) = \sum_{n=0}^{M-1} f(n) f(n-m)$
### Fitting. Analysis of the residuals. Non “ideal” case

#### Ag β decay

**Model:** ExpDec2

**Equation:**

\[ y = A_1 \exp(-\lambda_1 t) + A_2 \exp(-\lambda_2 t) + y_0 \]

**Reduced Chi-Sqr:** 100.10041

**Adj. R-Square:** 0.99181

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Value</th>
<th>Standard Error</th>
</tr>
</thead>
<tbody>
<tr>
<td>( y_0 )</td>
<td>5.18284</td>
<td>1.99542</td>
</tr>
<tr>
<td>A1</td>
<td>130.85655</td>
<td>20.27379</td>
</tr>
<tr>
<td>( t_1 )</td>
<td>145.8949</td>
<td>21.82649</td>
</tr>
<tr>
<td>A2</td>
<td>792.62197</td>
<td>19.21953</td>
</tr>
<tr>
<td>( t_2 )</td>
<td>27.83839</td>
<td>1.30697</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>Clear experiment</th>
<th>Data + “noise”</th>
</tr>
</thead>
<tbody>
<tr>
<td>( t_1(s) )</td>
<td>177.76</td>
<td>145.89</td>
</tr>
<tr>
<td>( t_2(s) )</td>
<td>30.32</td>
<td>27.94</td>
</tr>
</tbody>
</table>

---

![Graph showing count over time for Ag β decay experiment data and residuals.](image-url)
Fitting. Analysis of the residuals. Non “ideal” case

Ag $\beta$ decay

Histogram does not follow the normal distribution and there is frequency of 0.333 is present in spectrum
Fitting. Analysis of the residuals. Non “ideal” case

Conclusion: fitting function should be modified by adding an additional term:

\[ y(t) = y_0 + A_1 \exp\left(\frac{-t}{t_1}\right) + A_2 \exp\left(\frac{-t}{t_2}\right) + A_3 \sin(\omega t + \theta) \]
Fitting. Analysis of the residuals. Non “ideal” case

<table>
<thead>
<tr>
<th></th>
<th>Clear experiment</th>
<th>Data + noise</th>
<th>Modified fitting</th>
</tr>
</thead>
<tbody>
<tr>
<td>$t_1(s)$</td>
<td>177.76</td>
<td>145.89</td>
<td>172.79</td>
</tr>
<tr>
<td>$t_2(s)$</td>
<td>30.32</td>
<td>27.94</td>
<td>30.17</td>
</tr>
</tbody>
</table>

FFT

Autocorrelation
Error propagation

\[ y = f(x_1, x_2 \ldots x_n) \]

\[
\Delta f(x_i, \Delta x_i) = \sqrt{\sum_{i=1}^{n} \left( \frac{\partial f}{\partial x_i} \right)^2} \cdot \Delta x_i^2
\]
Derive resonance frequency $f$ from measured inductance $L \pm \Delta L$ and capacitance $C \pm \Delta C$.

$$f (L, C) = \frac{1}{2\pi} \sqrt{\frac{1}{LC}}$$

$L_1 = 10 \pm 1\text{mH}, \quad C_1 = 10 \pm 2\mu\text{F}$

$$\Delta f (L, C, \Delta L, \Delta C) = \sqrt{\left[ \frac{\partial f}{\partial L} \right]^2 \Delta L^2 + \left[ \frac{\partial f}{\partial C} \right]^2 \Delta C^2}$$

$$\frac{\partial f}{\partial L} = \frac{-1}{4\pi} C^{-\frac{1}{2}} L^{-\frac{3}{2}};$$

$$\frac{\partial f}{\partial C} = \frac{-1}{4\pi} L^{-\frac{1}{2}} C^{-\frac{3}{2}}$$

**Results:**

$f(L_1, C_1) = 503.29212104487\text{Hz}$

$\Delta f = 56.26977\text{Hz}$

$f(L_1, C_1) = 503.3\pm56.3\text{Hz}$
In general we could expect both components of errors

\[ Q_{\text{meas}} = Q_{\text{true}} + e_s + e_r \]

\( e_s \) - systematic error comes from uncertainties of plates separation distance, applied DC voltage, ambient temperature etc.

\[ V = V_{DC} \pm \Delta V, \ d = d_0 \pm \Delta d \ ... \]

\( e_r \) - random errors are related to uncertainty of the knowledge of the actual \( t_g \) and \( t_{\text{rise}} \).

Uncertainty of time of crossing the marker line. It is random.

\[
Q = F \cdot S \cdot T = \left( \frac{1}{f_c^{3/2}} \right) \frac{9 \pi d}{V} \sqrt{\frac{2 \eta^3 x^3}{g \rho}} \frac{1}{\sqrt{t_g}} \left( \frac{1}{t_g} + \frac{1}{t_{\text{rise}}} \right)
\]

\[
\Delta Q = \sqrt{(S \cdot T)^2} \Delta F^2 + (F \cdot T)^2 \Delta S^2 + (F \cdot S)^2 \Delta T^2
\]

\[
T = \frac{1}{\sqrt{t_g}} \left( \frac{1}{t_g} + \frac{1}{t_{\text{rise}}} \right)
\]

\[
\Delta T = \sqrt{\left( \frac{3/2}{t_g^{5/2}} + \frac{1/2}{t_g^{3/2}} \frac{1}{t_{\text{rise}}} \right)^2 \Delta t_g^2 + \left( \frac{1}{t_g^{1/2}} \frac{1}{t_{\text{rise}}^2} \right)^2 \Delta t_{\text{rise}}^2}
\]
Step 1. Collect your data + parameters of the experiment in:

\[\text{Path: Phyap|portal\PHYCS401\Common\Origin templates\Oil drop experiment\Section L1.opj}\]

Use different columns for each student or team. This Origin project is for data collecting only but not for data analysis. For data analysis you have to copy these data and experiment parameters obtained by different students/team and paste it in one in your personal Origin project.
Appendix #1. Analyzing of the statistical data.

Step 2. Working on personal Origin project

Make a copy of the Millikan1 project to your personal folder and open it.

Paste these 5 parameters and raw data from Section L1-L4.opj projects.

Calculate manually the actual air viscosity.

Prepare equations calculations of data in next columns (Set column values...). Switch Recalculate in Auto mode.

#### Table

<table>
<thead>
<tr>
<th>Long Name</th>
<th>Parameter names</th>
<th>Parameter label</th>
<th>Par</th>
<th>tr</th>
<th>rc</th>
<th>tau_g</th>
<th>F</th>
</tr>
</thead>
<tbody>
<tr>
<td>Units</td>
<td>s</td>
<td>s</td>
<td>m</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Comments</td>
<td>your data</td>
<td>your data</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

- Actual air viscosity
- Voltage across the plates (V)
- Air temperature (OC)
- Voltage across the plates (V)
- Voltage across the plates (V)
Appendix #1. Analyzing of the statistical data.

Step 3. Histogram graph

First use the data from the column with drop charges and plot the histogram.

Millikan oil drop experiment
Step 4. Histogram. Bin size

Origin will automatically but not optimally adjust the bin size \( h \). In this page figure \( h=0.5 \). There are several theoretical approaches how to find the optimal bin size.

\[
h = \frac{3.5\sigma}{n^{1/3}}
\]

\( \sigma \) is the sample standard deviation and \( n \) is total number of observation. For presented in Fig.1 results good value of \( h \sim 0.1 \)
Appendix #1. Analyzing of the statistical data.

**Step 4. Histogram. Bin size**

To change the bin size click on graph and unplug the “**Automatic Binning**” option.

Bin size in this histogram is 0.1
Appendix #1. Analyzing of the statistical data.

Step 4. Multipeak Gaussian fitting

Millikan oil drop experiment

To do this you have to add an extra plot to the graph

Counts vs. Bin Center
Appendix #1. Analyzing of the statistical data.

Step 4. Multipeak Gaussian fitting

This plot can be used for peak fitting.
Step 4. Multipeak Gaussian fitting

This plot can be used for peak fitting.

Final result for first two peaks:

\[
\frac{Q}{e} = 0.93 \pm 0.01
\]

\[
\frac{Q}{e} = 1.87 \pm 0.02
\]

This pretty close to \( e \) and \( 2e \)

Here \( w = 2\sigma \) and error of the mean = \( \frac{\sigma}{\sqrt{N}} \)