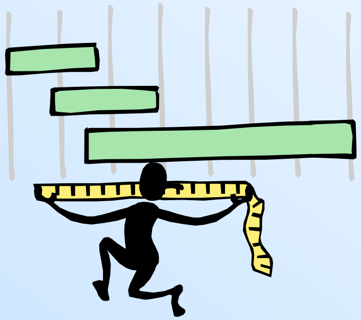
A Review of Uncertainty Analysis

MANE-4430

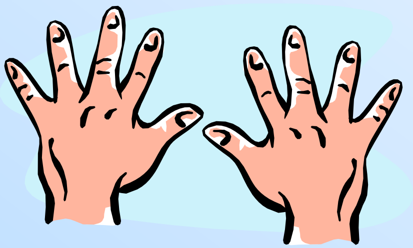
LINEAR ACCELERATOR LAB

Error Analysis

- In the laboratory we measure physical quantities.
- All measurements are subject to some uncertainties.
- Error analysis is the study and evaluation of these uncertainties.
- When mathematically manipulating measured quantities, a proper manipulation is required for the uncertainties.



Errors in Measurements

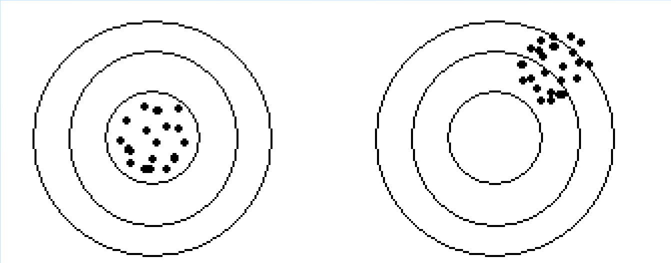
- Measure length with ruler. 
- Measure voltage with digital multimeter. 
- Measure time. 
- Measure radiation decay (counting statistics).
- Stated instrumentation accuracy.
- Counting ?? 

Types of Uncertainties

- RANDOM – arising from a random effect.
  - Example: radioactive nuclear decay.

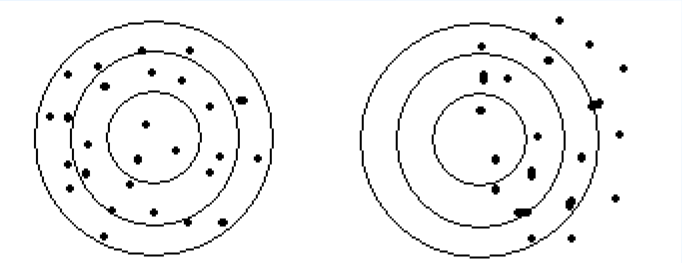
- SYSTEMATIC – arising from a systematic effect.
  - Example: instrument calibration error.

Examples



Random: Small
Systematic: Small

Random: Small
Systematic: Large

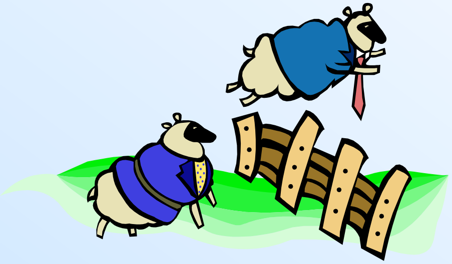


Random: Large
Systematic: Small

Random: Large
Systematic: Large

Reporting Uncertainties

- (Measured value of x)=$x \pm \delta x$
  - Example V=2.5±0.2
- Round error to one significant digit.
  - ~~g=9.82±0.02385 m/s^2~~ → g=9.82±0.02 m/s^2
- The last significant digit of the quantity should be of the same order of the uncertainty.
  - ~~I=4.35±0.2 A~~ → I=4.4±0.2 A



Counting Statistics I

- For a process with very small success probability $p<<1$, if we carry n experiments, the distribution having x successes is Binomial. It can be approximated by the Poisson distribution.

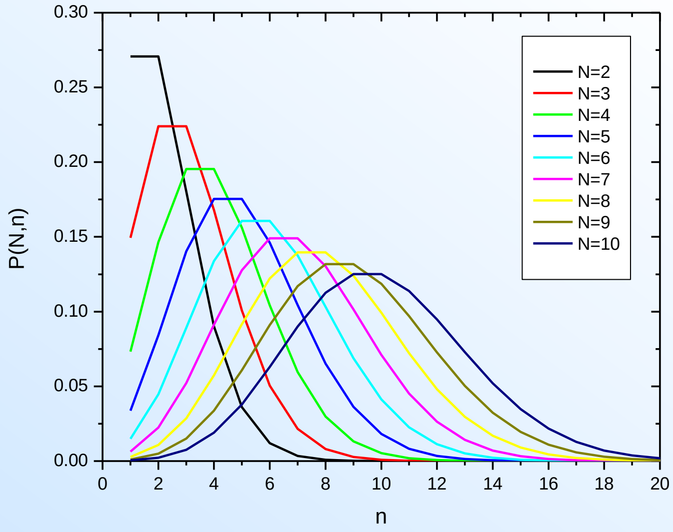
$$p(x) = \frac{(pn)^x e^{-pn}}{x!}$$

- The average and variance of this distribution is pn.
- In nuclear decay, large number of nuclei make up a sample or numbers of tries (n) but only a relatively small fraction of them give rise to a success event (small p).
- In a counting experiment we record the number of counts, n in a given counting time t. The distribution of n is Poisson:

$$\boxed{p(n) = \frac{N^n e^{-N}}{n!}}$$

- Where N is the expected counts (the mean) and uncertainty $\sigma = \sqrt{N}$

Counting Statistics II



- For large N the Poisson distribution can be approximated by the Gaussian distribution.
- **Example** : we make N experiment and record x_i counts in each. What is the average number of counts and expected error of the average?

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

$$\sigma = \frac{1}{N}\sqrt{\sum_{i=1}^{N}(\sqrt{x_i})^2} = \frac{1}{N}\sqrt{\sum_{i=1}^{N} x_i} = \frac{1}{N}\sqrt{N\bar{X}} = \sqrt{\frac{\bar{x}}{N}}$$

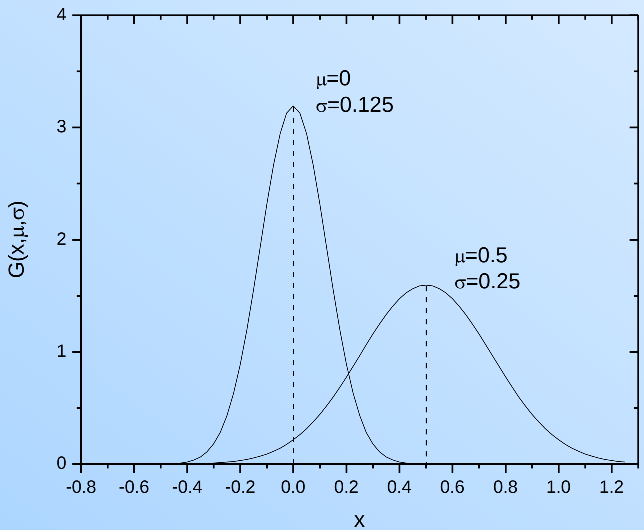
Measurement Distribution I

- In most cases the distribution of a measured quantity is Gaussian (or Normal when $\sigma = \sqrt{\bar{x}}$).

$$G(x,\mu,\sigma) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2}$$

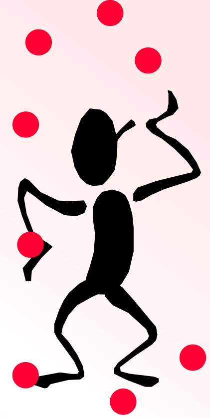
- Where μ is the average and σ is the standard deviation.

$$\mu = \int_{-\infty}^{\infty} G(x,\mu,\sigma) x dx \qquad \sigma^2 = \int_{-\infty}^{\infty} (x-\mu)^2 G(x,\mu,\sigma) dx$$



- If *x* is sampled from a Normal distribution then
- 68% of the samples will be between μ-σ to μ+σ.
- 95.5% between μ-2σ to μ+2σ.
- 99.7% between μ-3σ to μ+3σ.

Error Propagation I



- Consider a function $q(x_i,y_i)$ $I=1,\ldots N.$
- The first order Taylor series expansion of $q(x_i,y_i)$ at the point $(\bar{x},\bar{y})$

$$q_i = q(x_i, y_i) \approx q(\bar{x},\bar{y}) + \left.\frac{\partial q}{\partial x}\right|_{\bar{x},\bar{y}} (x_i - \bar{x}) + \left.\frac{\partial q}{\partial y}\right|_{\bar{x},\bar{y}} (y_i - \bar{y})$$

- We can calculate the mean of $q(x_i,y_i)$:

$$\bar{q} = \frac{1}{N}\sum_{i=1}^{N} q_i \approx \frac{1}{N}\sum_{i=1}^{N}\left[q(\bar{x},\bar{y}) + \left.\frac{\partial q}{\partial x}\right|_{\bar{x},\bar{y}} (x_i - \bar{x}) + \left.\frac{\partial q}{\partial y}\right|_{\bar{x},\bar{y}} (y_i - \bar{y})\right]$$

- Which can be written as:

$$\bar{q} = \frac{1}{N}\sum_{i=1}^{N} q_i \approx \frac{1}{N}\sum_{i=1}^{N} q(\bar{x},\bar{y}) + \left.\frac{\partial q}{\partial x}\right|_{\bar{x},\bar{y}} \frac{1}{N}\sum_{i=1}^{N}(x_i - \bar{x}) + \left.\frac{\partial q}{\partial y}\right|_{\bar{x},\bar{y}} \frac{1}{N}\sum_{i=1}^{N}(y_i - \bar{y})$$

- From the definition of the average $\sum_{i=1}^{N}(x_i - \bar{x}) = 0, \sum_{i=1}^{N}(y_i - \bar{y}) = 0$ and thus

$$\bar{q} = \frac{1}{N}\sum_{i=1}^{N} q(\bar{x},\bar{y}) = \frac{q(\bar{x},\bar{y})}{N}\sum_{i=1}^{N} 1 \Rightarrow \qquad \boxed{\bar{q} = q(\bar{x},\bar{y})}$$

Error Propagation II

- The variance of q is defined as : $\sigma_q^2 = \frac{1}{N}\sum_{i=1}^{N}(q_i - \bar{q})^2$
- Evaluating the variance we get:

$$\sigma_q^2 = \frac{1}{N}\sum_{i=1}^{N}\left[\frac{\partial q}{\partial x}(x_i - \bar{x}) + \frac{\partial q}{\partial y}(y_i - \bar{y})\right]^2 = \left(\frac{\partial q}{\partial x}\right)^2 \frac{1}{N}\sum_{i=1}^{N}(x_i - \bar{x})^2 + \left(\frac{\partial q}{\partial y}\right)^2 \frac{1}{N}\sum_{i=1}^{N}(y_i - \bar{y})^2$$

$$+ 2\frac{\partial q}{\partial x}\frac{\partial q}{\partial y}\frac{1}{N}\sum_{i=1}^{N}(x_i - \bar{x})(y_i - \bar{y})$$

We define the covariance:

$$\sigma_{xy} = \frac{1}{N}\sum_{i=1}^{N}(x_i - \bar{x})(y_i - \bar{y})$$

- And finally the standard deviation σ_q is given by:

$$\sigma_q^2 = \left(\frac{\partial q}{\partial x}\right)^2 \sigma_x^2 + \left(\frac{\partial q}{\partial y}\right)^2 \sigma_y^2 + 2\frac{\partial q}{\partial x}\frac{\partial q}{\partial y}\sigma_{xy}$$

Error Propagation – Examples I

- In many cases we can assume that the variables are independent. For a function with *n* variables $q(x_1,x_2,\ldots.x_n)$ the variance is given by:

$$\sigma_q^2 = \sum_{i=1}^{n} \left(\frac{\partial q}{\partial x_i} \right)^2 \sigma_i^2$$

- Using this equation, we can derive simple helpful relations for the propagation of errors:

- Addition and subtraction: $u = x \pm y \quad \frac{\partial u}{\partial x} = 1 \quad \frac{\partial u}{\partial y} = 1 \Rightarrow \Delta u = \sqrt{\Delta x^2 + \Delta y^2}$

- Multiplication by a constant: $u = Ax \quad \frac{\partial u}{\partial x} = A \quad \Rightarrow \Delta u = A\Delta x$

- Multiplication or division: $u = xy$ or $u = \frac{x}{y} \qquad \left(\frac{\Delta u}{u}\right)^2 = \left(\frac{\Delta x}{x}\right)^2 + \left(\frac{\Delta y}{y}\right)^2$

Error Propagation – Example II

- we measured V=1.51±0.02 V across a resistor with R=900±5 Ω What is the current I.

$$I = \frac{V}{R} = \frac{1.51}{900} = 1.67778 \times 10^{-3} \text{ A}$$

- Use the error propagation for division, the fractional error of I is:

$$\frac{\Delta I}{I} = \sqrt{\left(\frac{\Delta R}{R}\right)^2 + \left(\frac{\Delta V}{V}\right)^2} = \sqrt{\left(\frac{5}{900}\right)^2 + \left(\frac{0.02}{1.51}\right)^2} = 0.01436$$

- The error in I is then $\Delta I = I \frac{\Delta I}{I} = 0.01436 \times 1.67778 \times 10^{-3} = 0.024 \times 10^{-3} \text{ A}$
- We report: I=1.68±0.02 mA.

Covariance

- A covariance value that is different from zero indicates that data are correlated. Here is an example.

$$\sigma_{xy} = \frac{1}{N}\sum_{i=1}^{N}(x_i - \bar{x})(y_i - \bar{y})$$

ORIGIN := 1

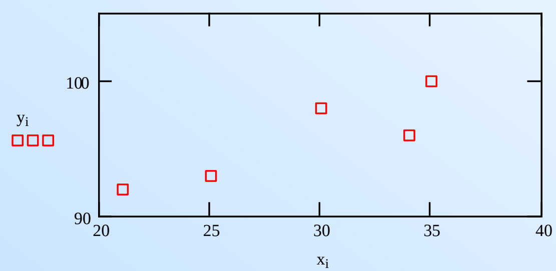
$$x := \begin{pmatrix} 25 \\ 34 \\ 35 \\ 30 \\ 21 \end{pmatrix} \qquad y := \begin{pmatrix} 93 \\ 96 \\ 100 \\ 98 \\ 92 \end{pmatrix} \qquad N := \mathrm{length}(x)$$

$x_{av} := \mathrm{mean}(x)$ $\quad x_{av} = 29$ $\quad y_{av} := \mathrm{mean}(y)$ $\quad y_{av} = 95.8$

$\sigma_x := \sqrt{\mathrm{var}(x)}$ $\quad \sigma_x = 5.329$ $\quad \sigma_y := \sqrt{\mathrm{var}(y)}$ $\quad \sigma_y = 2.993$

$$\sigma_{xy} := \frac{1}{N}\cdot\sum_{i=1}^{N}(x_i - x_{av})(y_i - y_{av}) \qquad \sigma_{xy} = 14$$

$i := 1..N$



Assume w e w ant to find the average sum <z>=<x>+<y>

$z_{av} := x_{av} + y_{av}$ $\quad z_{av} = 124.8$

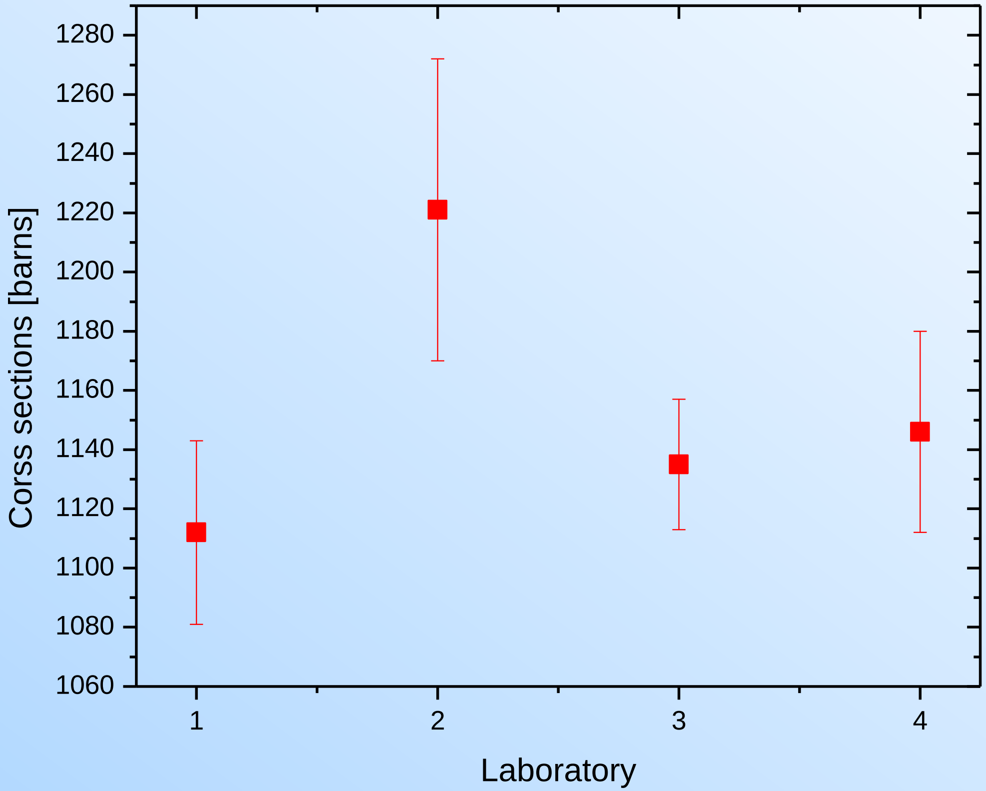
without covariance **with covariance**

$\sigma z1 := \sqrt{\sigma_x^2 + \sigma_y^2}$ $\quad \sigma z2 := \sqrt{\sigma_x^2 + \sigma_y^2 + 2\cdot\sigma_{xy}}$

$\sigma z1 = 6.112$ $\quad \sigma z2 = 8.085$

Discrepant measurements

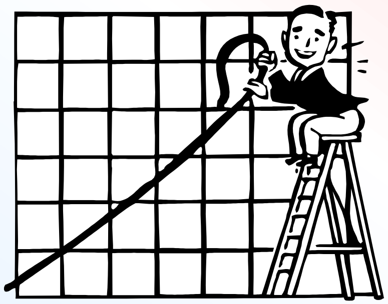
- 4 laboratories measured the absorption cross section of the same isotope.
- Which value should I use ?



Average and Variance

- Given N measurements of a quantity x_i, we can estimate the mean of the distribution of x by calculating the average:

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



- The estimate for the variance of the distribution is:

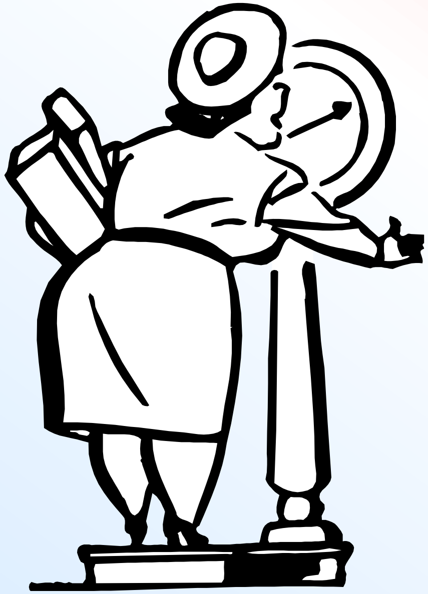
$$\sigma^2 = \frac{1}{N-1}\sum_{i=1}^{N}(x_i - \bar{x})^2$$

- The estimated standard deviation is σ.
- The more samples we have (larger N), the average $\bar{x}$ will get closer to the real average $\bar{X}$ of the distribution.

Weighted Average

- When we measure the quantity x_i with an associated error σ_i, then the *best estimate of the of that quantity* is calculated by the weighted average:

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



- Where the weight is taken as $w_i = \frac{1}{\sigma_i^2}$
- The error in that estimated average is be given by:

$$\sigma_w = \frac{1}{\sqrt{\sum_{i=1}^{N} w_i}}$$

- *This calculation gives more weight to measurements with reported small errors.*

Average and Variance - Example

ORIGIN := 1

Following are results of the same measurement from several students

$$x := \begin{pmatrix} 13 \\ 12 \\ 16 \\ 11.5 \end{pmatrix} \qquad \sigma := \begin{pmatrix} 0.5 \\ 0.4 \\ 2 \\ 0.3 \end{pmatrix} \qquad N := \text{length}(x)$$

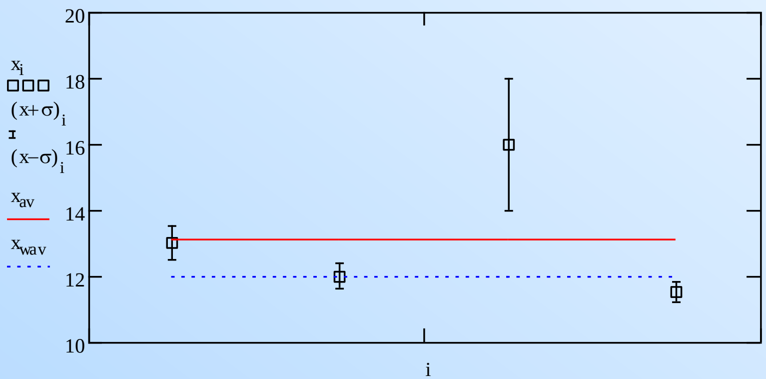
Non Weighted

$$x_{av} := \frac{1}{N} \cdot \sum_{i=1}^{N} x_i \qquad x_{av} = 13.125 \qquad std := \sqrt{\frac{1}{N} \cdot \sum_{i=1}^{4} \left(x_i - x_{av}\right)^2} \qquad std = 1.746$$

Weighted

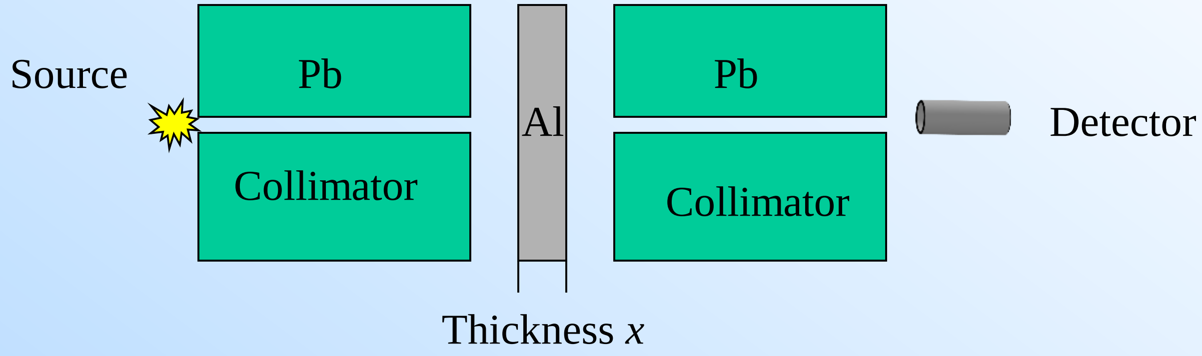
$$i := 1..N \qquad w_i := \frac{1}{(\sigma_i)^2}$$

$$x_{wav} := \frac{\sum_{i=1}^{N} (w_i \cdot x_i)}{\sum_{i=1}^{N} w_i} \qquad x_{wav} = 11.974 \qquad \sigma_{wav} := \frac{1}{\sqrt{\sum_{i=1}^{N} w_i}} \qquad \sigma_{wav} = 0.215$$



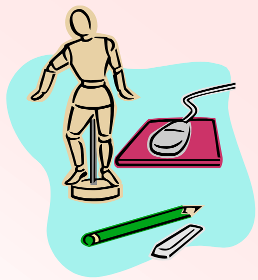
Modeling of Data-I

- In many cases we have some theoretical background on the physical behavior of the phenomena we are measuring.
- In these cases we can try to check if the experiment agrees with the theory and also extract parameters from it.
- For example we would like to measure the attenuation of gamma rays through a slab of Al. We setup the following experiment:



- We repeat the experiment *N* times (*N*>3) for several thickness x_i of Aluminum
- We have no background (or corrected for it).

Modeling of Data-II



- If we count for sufficient time we except:

$$C_i = I_0 e^{-\mu x_i}$$

- Where C_i are the counts for sample i, μ is the attenuation coefficient in units of cm^{-1} and x_i is the sample thickness.
- We take the log of both sides of the the equation we get:

$$\log(C_i) = \log(I_0) - \mu x_i$$

- Define $y_i = \log(C_i) \quad \mathrm{b} = -\mu \quad a = \log(I_0)$
- The equation can be rewritten as:

$$y_i = bx_i + a$$

- Find a procedure that can use all our measurements of y_i to find a and b that best fit the data.
- Knowing the counting error in C_i, we will try to estimate the error in a and b.

Least-Squares Fitting I

- Given a series of N measurement of x_i and $y_i \pm \sigma_{i.}$, Fit the model $y_i = a + bx_i$ to the data.
- We would like find a and b that will minimize the expression.

$$\chi^2 = \sum_{i=1}^{N} \left(\frac{y_i - a - b_i x_i}{\sigma_i}\right)^2$$

- To do that we take the first derivative with respect to a and b and set the derivatives equal to zero:

$$\frac{\partial \chi^2}{\partial a} = -2 \sum_{i=1}^{N} \frac{y_i - a - bx_i}{\sigma_i} = 0$$

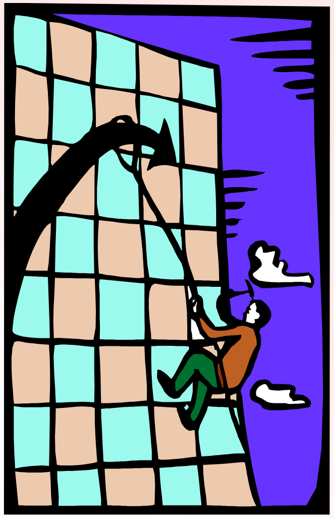
$$\frac{\partial \chi^2}{\partial b} = -2 \sum_{i=1}^{N} \frac{y_i - a - bx_i}{\sigma_i} x_i = 0$$

- We can define:

$$S \equiv \sum_{i=1}^{N} \frac{1}{\sigma_i^2} \qquad S_x \equiv \sum_{i=1}^{N} \frac{x_i}{\sigma_i^2} \qquad S_y \equiv \sum_{i=1}^{N} \frac{y_i}{\sigma_i^2}$$

$$S_{xx} \equiv \sum_{i=1}^{N} \frac{x_i^2}{\sigma_i^2} \qquad S_{xy} \equiv \sum_{i=1}^{N} \frac{x_i y_i}{\sigma_i^2}$$

Least-Squares Fitting II



- We can then rewrite the equations as:

$$aS + bS_x = S_y$$
$$aS_x + bS_{xx} = S_{xy}$$

- We solve the equations for *a* and *b*:

$$a = \frac{S_{xx}S_y - S_x S_{xy}}{\Delta} \qquad b = \frac{SS_{xy} - S_x S_y}{\Delta}$$
$$\Delta = SS_{xx} - (S_x)^2$$

- The error in *a* and *b* can be estimated by propagating the errors in the above equations (independent case) , the result is:

$$\sigma_a = \sqrt{\frac{S_{xx}}{\Delta}} \qquad \sigma_b = \sqrt{\frac{S}{\Delta}}$$

Least-Squares Fitting - Example

$$x = \begin{pmatrix} 2 \\ 4 \\ 6 \\ 8 \\ 10 \end{pmatrix} \quad y1 = \begin{pmatrix} 431.219 \\ 184.518 \\ 81.06 \\ 31.792 \\ 9.71 \end{pmatrix} \quad \sigma1 := \sqrt{y1} \quad N := \text{length}(x)$$

Transform the data

$$y := \ln(y1) \quad \sigma := \overrightarrow{\frac{\sigma1}{y1}}$$

Remember : $y = I_0 \exp(-\mu x) \Rightarrow \ln(y) = \ln(y_0) - \mu x$

Fit the data

$$S := \sum_{i=1}^{N} \frac{1}{(\sigma_i)^2} \quad Sx := \sum_{i=1}^{N} \frac{x_i}{(\sigma_i)^2} \quad Sy := \sum_{i=1}^{N} \frac{y_i}{(\sigma_i)^2} \quad Sxx := \sum_{i=1}^{N} \frac{(x_i)^2}{(\sigma_i)^2} \quad Sxy := \sum_{i=1}^{N} \frac{x_i \cdot y_i}{(\sigma_i)^2}$$

$$\Delta := S \cdot Sxx - Sx^2 \quad a := \frac{Sxx \cdot Sy - Sx \cdot Sxy}{\Delta} \quad b := \frac{S \cdot Sxy - Sx \cdot Sy}{\Delta} \quad \sigma_a := \sqrt{\frac{Sxx}{\Delta}} \quad \sigma_b := \sqrt{\frac{S}{\Delta}}$$

Transforming the results of the fit

$$\mu := -b \quad \sigma_\mu := \sigma_b \quad I_0 := \exp(a) \quad \sigma_{I0} := \exp(a) \cdot \sigma_a$$

$\mu = 0.44$ $\quad \sigma_\mu = 0.02$ $\quad I_0 = 1041$ $\quad \sigma_{I0} = 78$

$$j := 1 .. 100 \quad xx_j := 0.1 \cdot j \quad I_j := I_0 \cdot \exp(-\mu \cdot xx_j)$$