Lecture 7
Some Advanced Topics using Propagation of Errors and Least Squares Fitting

Error on the mean (review from Lecture 4)

- Question: If we have a set of measurements of the same quantity:
  $x_1 \pm \sigma_1 \;\; x_2 \pm \sigma_2 \ldots \; x_n \pm \sigma_n$
  - What's the best way to combine these measurements?
  - How to calculate the variance once we combine the measurements?
  - Assuming Gaussian statistics, the Maximum Likelihood Methods combine the measurements as:

$$x = \frac{\sum_{i=1}^{n} x_i/\sigma_i^2}{\sum_{i=1}^{n} 1/\sigma_i^2}$$

weighted average

  - If all the variances ($\sigma_1^2 = \sigma_2^2 = \ldots \sigma_n^2$) are the same:

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

unweighted average

  - The variance of the weighted average can be calculated using propagation of errors:

$$\sigma_x^2 = \sum_{i=1}^{n}\left[\frac{\partial}{\partial x_i}x\right]^2 \sigma_i^2 = \sum_{i=1}^{n}\frac{1/\sigma_i^4}{\left[\sum_{i=1}^{n} 1/\sigma_i^2\right]^2}\sigma_i^2 = \frac{1}{\left[\sum_{i=1}^{n} 1/\sigma_i^2\right]^2}\sum_{i=1}^{n} 1/\sigma_i^2$$

$$\sigma_x^2 = \frac{1}{\sum_{i=1}^{n} 1/\sigma_i^2}$$

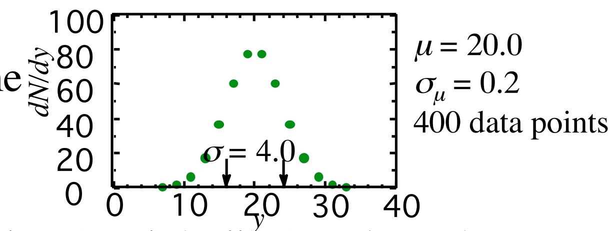
σ_x is the error in the weighted mean

- If all the variances are the same:

$$\sigma_{\bar{x}}^2 = 1/\sum_{i=1}^{n} 1/\sigma_i^2 = 1/[n/\sigma^2] = \frac{\sigma^2}{n}$$

Lecture 4

  - ☞ The error in the mean ($\sigma_{\bar{x}}$) gets smaller as the number of measurements (n) increases.
  - Don't confuse the error in the mean ($\sigma_{\bar{x}}$) with the standard deviation of the distribution (σ)!
  - If we make more measurements
    - ☞ the standard deviation (σ) of the distribution remains the same
    - ☞ the error in the mean ($\sigma_{\bar{x}}$) decreases



More on Least Squares Fit (LSQF)

- In Lec 5, we discussed how we can fit our data points to a linear function (straight line) and get the "best" estimate of the slope and intercept. However, we did not discuss two important issues:
  - How to estimate the uncertainties on our slope and intercept obtained from a LSQF?
  - How to apply the LSQF when we have a non-linear function?
- Estimation of Errors from a LSQF
  - Assume we have data points that lie on a straight line:

    $y = \alpha + \beta x$

    - Assume we have n measurements of x's and y's.
    - For simplicity, assume that each y measurement has the same error σ.
    - Assume that x is known much more accurately than y.
      - ☞ ignore any uncertainty associated with x.
    - Previously we showed that the solution for the intercept α and slope β is:

$$\alpha = \frac{\sum_{i=1}^{n} y_i \sum_{i=1}^{n} x_i^2 - \sum_{i=1}^{n} x_i y_i \sum_{i=1}^{n} x_i}{n\sum_{i=1}^{n} x_i^2 - (\sum_{i=1}^{n} x_i)^2} \quad \text{and} \quad \beta = \frac{n\sum_{i=1}^{n} x_i y_i - \sum_{i=1}^{n} x_i \sum_{i=1}^{n} y_i}{n\sum_{i=1}^{n} x_i^2 - (\sum_{i=1}^{n} x_i)^2}$$

- Since α and β are functions of the measurements (y_i's)
  - use the Propagation of Errors technique to estimate σ_α and σ_β.

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

  - ★ Assumed that each measurement is independent of each other:

$$\sigma_Q^2 = \sigma_x^2\left(\frac{\partial Q}{\partial x}\right)^2 + \sigma_y^2\left(\frac{\partial Q}{\partial y}\right)^2$$

$$\sigma_\alpha^2 = \sum_{i=1}^{n}\sigma_{y_i}^2\left(\frac{\partial \alpha}{\partial y_i}\right)^2 = \sigma^2\sum_{i=1}^{n}\left(\frac{\partial \alpha}{\partial y_i}\right)^2$$

$$\frac{\partial \alpha}{\partial y_i} = \frac{\partial}{\partial y_i}\frac{\sum_{i=1}^{n} y_i \sum_{j=1}^{n} x_j^2 - \sum_{i=1}^{n} x_i y_i \sum_{j=1}^{n} x_j}{n\sum_{i=1}^{n} x_i^2 - (\sum_{i=1}^{n} x_i)^2} = \frac{\sum_{j=1}^{n} x_j^2 - x_i \sum_{j=1}^{n} x_j}{n\sum_{i=1}^{n} x_i^2 - (\sum_{i=1}^{n} x_i)^2}$$

$$\sigma_\alpha^2 = \sigma^2\sum_{i=1}^{n}\left(\frac{\sum_{j=1}^{n} x_j^2 - x_i \sum_{j=1}^{n} x_j}{n\sum_{i=1}^{n} x_i^2 - (\sum_{i=1}^{n} x_i)^2}\right)^2 = \sigma^2\sum_{i=1}^{n}\left(\frac{(\sum_{j=1}^{n} x_j^2)^2 + x_i^2(\sum_{j=1}^{n} x_j)^2 - 2x_i\sum_{j=1}^{n} x_j\sum_{j=1}^{n} x_j^2}{(n\sum_{i=1}^{n} x_i^2 - (\sum_{i=1}^{n} x_i)^2)^2}\right)$$

$$\sigma_\alpha^2 = \sigma^2 \frac{n(\sum_{j=1}^n x_j^2)^2 + \sum_{i=1}^n x_i^2 (\sum_{j=1}^n x_j)^2 - 2(\sum_{j=1}^n x_j)^2 \sum_{j=1}^n x_j^2}{(n\sum_{i=1}^n x_i^2 - (\sum_{i=1}^n x_i)^2)^2} = \sigma^2 \frac{n(\sum_{j=1}^n x_j^2)^2 - \sum_{i=1}^n x_i^2 (\sum_{j=1}^n x_j)^2}{(n\sum_{i=1}^n x_i^2 - (\sum_{i=1}^n x_i)^2)^2}$$

$$= \sigma^2 \sum_{j=1}^n x_j^2 \frac{n\sum_{j=1}^n x_j^2 - (\sum_{j=1}^n x_j)^2}{(n\sum_{i=1}^n x_i^2 - (\sum_{i=1}^n x_i)^2)^2}$$

$$\sigma_\alpha^2 = \sigma^2 \frac{\sum_{j=1}^n x_j^2}{n\sum_{i=1}^n x_i^2 - (\sum_{i=1}^n x_i)^2}$$

variance in the intercept

★ We can find the variance in the slope (β) using exactly the same procedure:

$$\sigma_\beta^2 = \sum_{i=1}^n \sigma_{y_i}^2 \left(\frac{\partial \beta}{\partial y_i}\right)^2 = \sigma^2 \sum_{i=1}^n \left(\frac{\partial \beta}{\partial y_i}\right)^2 = \sigma^2 \sum_{i=1}^n \left(\frac{\partial}{\partial y_i} \frac{n\sum_{i=1}^n x_i y_i - \sum_{i=1}^n x_i \sum_{i=1}^n y_i}{n\sum_{i=1}^n x_i^2 - (\sum_{i=1}^n x_i)^2}\right)^2 = \sigma^2 \sum_{i=1}^n \left(\frac{n x_i - \sum_{j=1}^n x_j}{n\sum_{i=1}^n x_i^2 - (\sum_{i=1}^n x_i)^2}\right)^2$$

$$= \sigma^2 \frac{n^2 \sum_{j=1}^n x_j^2 + n(\sum_{j=1}^n x_j)^2 - 2n\sum_{i=1}^n x_i \sum_{j=1}^n x_j}{(n\sum_{i=1}^n x_i^2 - (\sum_{i=1}^n x_i)^2)^2} = \sigma^2 \frac{n^2 \sum_{j=1}^n x_j^2 - n(\sum_{j=1}^n x_j)^2}{(n\sum_{i=1}^n x_i^2 - (\sum_{i=1}^n x_i)^2)^2}$$

$$\sigma_\beta^2 = \frac{n\sigma^2}{n\sum_{i=1}^{n} x_i^2 - (\sum_{i=1}^{n} x_i)^2}$$

variance in the slope

- If we don't know the true value of σ,
  - ☞ estimate variance using the spread between the measurements (y_i's) and the fitted values of y:

$$\sigma^2 \approx \frac{1}{n-2}\sum_{i=1}^{n}(y_i - y_i^{fit})^2 = \frac{1}{n-2}\sum_{i=1}^{n}(y_i - \alpha - \beta x_i)^2$$

    - ★ $n - 2$ = number of degree of freedom

      = number of data points − number of parameters (α, β) extracted from the data
- If each y_i measurement has a different error σ_i:

$$\sigma_\alpha^2 = \frac{1}{D}\sum_{i=1}^{n}\frac{x_i^2}{\sigma_i^2}$$

$$\sigma_\beta^2 = \frac{1}{D}\sum_{i=1}^{n}\frac{1}{\sigma_i^2}$$

weighted slope and intercept

$$D = \sum_{i=1}^{n}\frac{1}{\sigma_i^2}\sum_{i=1}^{n}\frac{x_i^2}{\sigma_i^2} - (\sum_{i=1}^{n}\frac{x_i}{\sigma_i^2})^2$$

  - ★ The above expressions simplify to the "equal variance" case.
    - ❒ Don't forget to keep track of the "n's" when factoring out σ. For example:

$$\sum_{i=1}^{n}\frac{1}{\sigma_i^2} = \frac{n}{\sigma^2} \quad not \quad \frac{1}{\sigma^2}$$

- LSQF with non-linear functions:
  - For our purposes, a non-linear function is a function where one or more of the parameters that we are trying to determine (e.g. α, β from the straight line fit) is raised to a power other than 1.
    - Example: functions that are non-linear in the parameter τ:

$$y = A + x/\tau$$

$$y = A + x\tau^2$$

$$y = Ae^{-x/\tau}$$

      - These functions are linear in the parameters A.
  - The problem with most non-linear functions is that we cannot write down a solution for the parameters in a closed form using, for example, the techniques of linear algebra (i.e. matrices).
    - Usually non-linear problems are solved numerically using a computer.
    - Sometimes by a change of variable(s) we can turn a non-linear problem into a linear one.
      - Example: take the natural log of both sides of the above exponential equation:

$$\ln y = \ln A - x/\tau = C - Dx$$

        - A linear problem in the parameters C and D!
        - In fact its just a straight line!
        - To measure the lifetime τ (Lab 6) we first fit for D and then transform D into τ.
  - Example: Decay of a radioactive substance. Fit the following data to find N_0 and τ:

$$N = N_0 e^{-t/\tau}$$

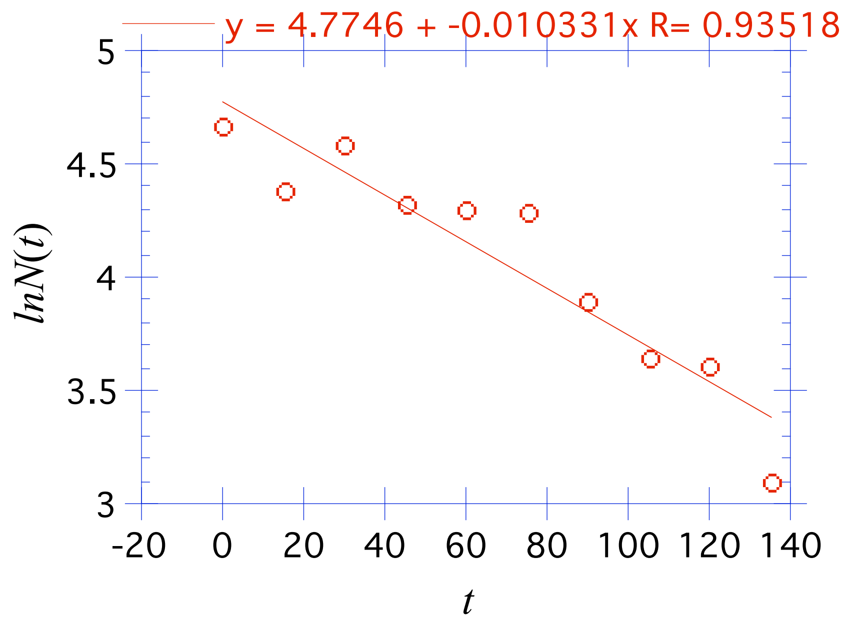
    - N represents the amount of the substance present at time t.
    - N_0 is the amount of the substance at the beginning of the experiment ($t = 0$).
    - τ is the lifetime of the substance.

i	1	2	3	4	5	6	7	8	9	10
t_i	0	15	30	45	60	75	90	105	120	135
N_i	106	80	98	75	74	73	49	38	37	22
$y_i = \ln N_i$	4.663	4.382	4.585	4.317	4.304	4.290	3.892	3.638	3.611	3.091

$$D = -\beta = -\frac{n\sum_{i=1}^{n} x_i y_i - \sum_{i=1}^{n} y_i \sum_{i=1}^{n} x_i}{n\sum_{i=1}^{n} x_i^2 - (\sum_{i=1}^{n} x_i)^2} = -\frac{10 \times 2560.41 - 40.773 \times 675}{10 \times 64125 - (675)^2} = 0.01033$$

$$\tau = 1/D = 96.80 \text{ sec}$$

- The intercept is given by: $C = 4.77 = \ln A$ or $A = 117.9$



- Example: Find the values A and τ taking into account the uncertainties in the data points.
  - The uncertainty in the number of radioactive decays is governed by Poisson statistics.
  - The number of counts N_i in a bin is assumed to be the average (μ) of a Poisson distribution:
    $\mu = N_i$ = Variance
  - The variance of y_i (= ln N_i) can be calculated using propagation of errors:
    $$\sigma_y^2 = \sigma_N^2 \left(\partial y / \partial N\right)^2 = (N)\left(\partial \ln N / \partial N\right)^2 = (N)\left(1/N\right)^2 = 1/N$$
  - The slope and intercept from a straight line fit that includes uncertainties in the data points:
    $$\alpha = \frac{\sum_{i=1}^{n} \frac{y_i}{\sigma_i^2} \sum_{i=1}^{n} \frac{x_i^2}{\sigma_i^2} - \sum_{i=1}^{n} \frac{x_i y_i}{\sigma_i^2} \sum_{i=1}^{n} \frac{x_i}{\sigma_i^2}}{\sum_{i=1}^{n} \frac{1}{\sigma_i^2} \sum_{i=1}^{n} \frac{x_i^2}{\sigma_i^2} - (\sum_{i=1}^{n} \frac{x_i}{\sigma_i^2})^2} \text{ and } \beta = \frac{\sum_{i=1}^{n} \frac{1}{\sigma_i^2} \sum_{i=1}^{n} \frac{x_i y_i}{\sigma_i^2} - \sum_{i=1}^{n} \frac{x_i}{\sigma_i^2} \sum_{i=1}^{n} \frac{y_i}{\sigma_i^2}}{\sum_{i=1}^{n} \frac{1}{\sigma_i^2} \sum_{i=1}^{n} \frac{x_i^2}{\sigma_i^2} - (\sum_{i=1}^{n} \frac{x_i}{\sigma_i^2})^2}$$

    Taylor P. 198 and Problem 8.9

    - ★ If all the σ's are the same then the above expressions are identical to the unweighted case.

    $\alpha = 4.725$ and $\beta = -0.00903$

    $\tau = -1/\beta = 1/0.00903 = 110.7$ sec
  - To calculate the error on the lifetime, we first must calculate the error on β:
    $$\sigma_\beta^2 = \frac{\sum_{i=1}^{n} \frac{1}{\sigma_i^2}}{\sum_{i=1}^{n} \frac{1}{\sigma_i^2} \sum_{i=1}^{n} \frac{x_i^2}{\sigma_i^2} - (\sum_{i=1}^{n} \frac{x_i}{\sigma_i^2})^2} = \frac{652}{652 \times 2684700 - (33240)^2} = 1.01 \times 10^{-6}$$
    $$\sigma_\tau^2 = \sigma_\beta^2 \left(\partial \tau / \partial \beta\right)^2 \Rightarrow \sigma_\tau = \sigma_\beta \left(1/\beta^2\right) = \frac{1.005 \times 10^{-3}}{(9.03 \times 10^{-3})^2} = 12.3$$

☞ The experimentally determined lifetime is
$\tau = 110.7 \pm 12.3$ sec.