Lecture 8
Some Advanced Topics

I) Error on the mean:

Question:

If we have a set of n measurements of the same quantity:

$$x_1 \pm \sigma_1 \ \ x_2 \pm \sigma_2 \ \ x_3 \pm \sigma_3 \ldots \ x_n \pm \sigma_n$$

what's the best way to combine these measurements and how do we calculate the variance once we combine the measurements?

We have already shown that, assuming Gaussian statistics, the Maximum Likelihood Method tells us to combine the measurements as follows:

$$\boxed{x = \frac{\sum_{i=1}^{n} x_i / \sigma_i^2}{\sum_{i=1}^{n} \sigma_i^{-2}}} \quad \text{weighted average}$$

We call this expression the "weighted average".

Note that if all the variances are the same ($\sigma_1 = \sigma_n$) then this reduces to the "ordinary" or "unweighted average":

$$\boxed{x = \frac{\sum_{i=1}^{n} x_i}{n}} \quad \text{unweighted average}$$

We would like to calculate the variance of the weighted average. We can do this using propagation of errors:

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

The derivative is evaluated as follows:

$$\frac{\partial}{\partial x_i} x = \frac{\sigma_i^{-2}}{\sum_{j=1}^{n} \sigma_j^{-2}}$$

Plugging this into the propagation of errors formula we have:

$$\sigma_x^2 = \sum_{i=1}^{n} \frac{\sigma_i^{-4}}{(\sum_{j=1}^{n} \sigma_j^{-2})^2} \sigma_i^2 = \frac{1}{(\sum_{j=1}^{n} \sigma_j^{-2})^2} \sum_{i=1}^{n} \sigma_i^{-2} = \frac{1}{\sum_{j=1}^{n} \sigma_j^{-2}}$$

or

$$\boxed{\sigma_x^2 = \frac{1}{\sum_{j=1}^{n} \sigma_j^{-2}}} \quad \text{error in the mean, weighted}$$

If all the variances are the same ($\sigma_1 = \sigma_n$) then the above expression simplifies to:

$$\boxed{\sigma_x^2 = \frac{\sigma^2}{n}} \quad \text{error in the mean, unweighted}$$

Thus the error in the mean gets smaller as the number of measurements increases.

Don't confuse the error in the mean with the variance of the distribution (σ)! As we make more measurements the variance of the distribution remains the same, however the error in the mean decreases.

II) More on Least Squares Fitting (see Taylor Chapt. 8)

1) *Introduction:*

Previously (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:

a) How do estimate the uncertainties on our slope and intercept obtained from a Least Squares Fit (LSQF)?

b) How do we apply LSQF when we have a non-linear function?

2) Estimation of Errors from a LSQF

Assume we have data points and believe that these points should lie on a straight line.

$$y = \alpha + \beta x$$

Here we have measured the y's (we have n measurements), and for the sake of simplicity let's assume that each y measurement has the same error σ. Also, assume that x is known much more accurately than y, we ignore any uncertainty associated with x. Previously we showed that the solution for α and β is:

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

We would like to calculate σ_α and σ_β.

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

Recall the Propagation of Errors formula:

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

We can neglect the third term in the above formula since we have assumed that each measurement is independent of each other. Thus we can simplify things and write:

$$\sigma_Q^2 = \sigma_x^2(\partial Q/\partial x)^2 + \sigma_y^2(\partial Q/\partial y)^2$$

First lets calculate σ_α:

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

We can factor out σ since we have assumed that each data point has the same uncertainty. We will consider the case where each of the measurements has a different uncertainty later.

The derivative of α is taken with respect to y_i, thus we must use a dummy index (j) to keep track of summations that are not affected by the derivative:

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

We can now plug the derivative into the propagation of errors formula to get:

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

We can simplify this by carrying out the summations:

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

Finally we find for the variance of the intercept of a straight line fit:

$$\boxed{\sigma_\alpha^2 = \frac{\sigma^2 \sum_{j=1}^{n} x_j^2}{n\sum_{i=1}^{n} x_i^2 - (\sum_{i=1}^{n} x_i)^2}} \quad \text{variance in the intercept}$$

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

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

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

Finally we obtain for the variance of the slope of a straight line fit:

$$\boxed{\sigma_\beta^2 = \frac{n\sigma^2}{n\sum_{j=1}^{n} x_j^2 - (\sum_{j}^{n} x_j)^2}} \quad \text{variance in the slope}$$

In the above derivation we assumed that all of the data points had the same uncertainty (σ). Suppose we don't know the true value of σ, is there anyway we can estimate it from our data and our fit? The answer of course is yes! We can estimate the variance using the spread between the measurements (y_i's) and the fitted values of y:

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

The above formula is similar to the definition of the variance, except that *n* (the number of data points) is replaced with *n - 2*. We use *n - 2* rather than *n* because we have determined two parameters (α, β) from the data. Thus *n - 2* represents the number of degree of freedom left in the problem.

Of course most of the time the individual uncertainties (σ_i) are not the same from measurement to measurement. In this case we most modify the above expressions, keeping track of the σ_i's. The results are:

$$\boxed{\sigma_\alpha^2 = \frac{1}{D} \sum_{i=1}^{n} \frac{x_i^2}{\sigma_i^2} \quad \text{and} \quad \sigma_\beta^2 = \frac{1}{D} \sum_{i=1}^{n} \frac{1}{\sigma_i^2} \quad \text{with } D \equiv \sum_{i=1}^{n} \frac{1}{\sigma_i^2} \sum_{i=1}^{n} \frac{x_i^2}{\sigma_i^2} - (\sum_{i=1}^{n} \frac{x_i}{\sigma_i^2})^2} \quad \text{variances, unequal errors}$$

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:

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

3) <u>LSQF and non-linear functions</u>:
First of all, what's a non-linear function? For our purposes a non-linear function is a function where the one or more of the parameters that we are trying determine (e.g. α, β from the straight line fit) is raised to a power other than 1. For example the following functions are non-linear in the parameter τ:

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

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

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

Note: these functions are linear in the parameters A. The problem with almost all 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, on a computer. However, sometimes by a change of variables we can turn a non-linear problem into a linear one. For example the function above with the exponential is non-linear in τ, however if we take the natural log of both sides of the equation we have:

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

Now this is a linear problem in the parameters C and D! In fact its just a straight line! So, if we are interested in τ (the lifetime in Lab 6) we could first fit for D and then transform D into τ.

<u>Example</u>: Using the "trick" discussed above fit the following data to an exponentially decreasing function, i.e. find A and τ:

$$N(t) = Ae^{-x/\tau}$$

This data could represent the decay of a radioactive substance. In this case N represents the amount of the substance present at a given time ("x"), A is the amount of the stuff at the beginning of the experiment (x = 0) and τ is the lifetime of the substance.

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

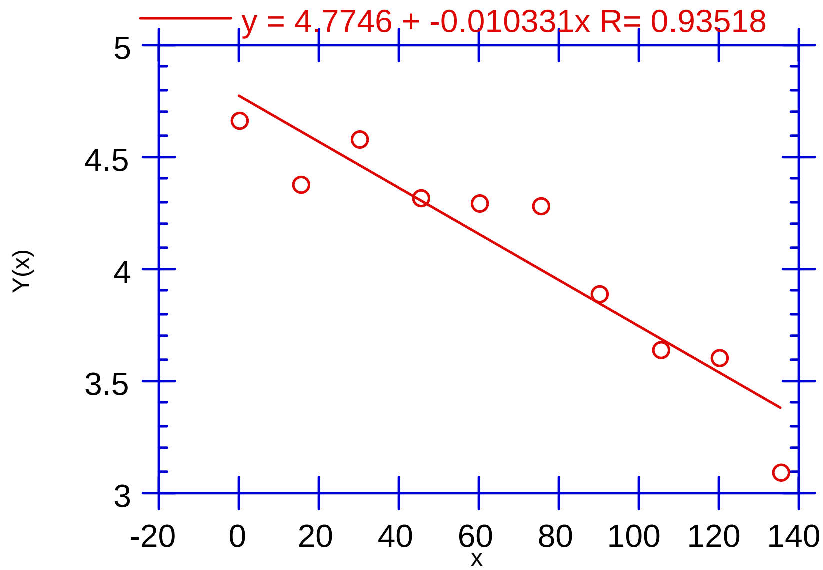
We have to evaluate several sums in order to calculate the slope:

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

The "lifetime" is just the inverse of D:

$$\tau = 1/D = 96.80$$

The figure below shows the data points and the result of the LSQF. The intercept is given by: $C = 4.77 = \ln A$ or $A = 117.9$



Example: Find the values A and τ taking account the uncertainties in the data points.

In the above example we did not "weight" the data with any σ's, we now want redo the problem taking into account the uncertainty in each measured point (σ_i). First what are the uncertainties in the measurements? For counting processes such as radioactive decay we assume that Poisson statistics apply. The recorded number of counts in a bin (N_i in our example) is assumed to be the average (μ) of a Poisson distribution for that bin. The variance of a Poisson distribution is just:

$$\text{Variance} = \mu = N_i$$

In order to take into account the individual uncertainties we must modify our expressions for the slope and intercept of a straight line:

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

Note: If all the σ's are the same then the above expressions are identical to the unweighted case. One sticky point remains before we can plug into the above equations. We know the variance of N_i, however for this problem we need the variance of $y_i (= \ln N_i)$. We must transform the variables using propagation of errors:

$$\sigma_y^2 = \sigma_N^2(\partial y / \partial N)^2 = (N)(\partial \ln N / \partial N)^2 = (N)(1/N)^2 = 1/N$$

Lots of sums to evaluate here! I used a MAC for the programming and got:

$$\alpha = \frac{2780.3 \times 2684700 - 132800 \times 33240}{652 \times 2684700 - (33240)^2} = 4.725$$

$$\beta = \frac{652 \times 132800 - 2780.3 \times 33240}{652 \times 2684700 - (33240)^2} = -0.00903$$

The lifetime (τ) is then:

$$\tau = 1/\beta = 1/0.00903 = 110.7 \text{ seconds}$$

Finally, lets calculate the error on β and the lifetime (τ):

$$\sigma_\beta^2 = \frac{\sum_{i=1}^{n}\frac{1}{\sigma_i^2}}{\sum_{i=1}^{n}\frac{1}{\sigma_i^2}\sum_{i=1}^{n}\frac{x_i^2}{\sigma_i^2} - (\sum_{i=1}^{n}\frac{x_i}{\sigma_i^2})^2} = \frac{652}{652 \times 2684700 - (33240)^2} = 1.01 \times 10^{-6}$$

To go from β to the lifetime, τ we again transform the variables:

$$\sigma_\tau^2 = \sigma_\beta^2 (\partial\tau / \partial\beta)^2 \Rightarrow \sigma_\tau = \sigma_\beta (1/\beta^2) = \frac{1.005 \times 10^{-3}}{(9.03 \times 10^{-3})^2} = 12.3$$

Thus the experimentally determined lifetime is:

$$\tau = 110.7 \pm 12.3 \text{ sec.}$$