Physics 433 - Lecture 7

- **Least square fit to a straight line**
  - Data consisting of \((x_i, y_i)\) pairs with \(x\) the independent and \(y\) the dependent variable related by \(y = A + Bx\)
  - **Method of least squares** that minimizes the differences
    \[
    \chi^2 = \frac{\left[y_i - y(x_i)\right]^2}{\sigma_i^2},
    \]
    which is called chi squared and is a measure of the “goodness of fit”
  - We assume that there is a “parent distribution” with the parent coefficients \(A_0\) and \(B_0\), where the “true” relationship between \(y\) and \(x\) is \(y_0(x) = A_0 + B_0x\) and further that each individual measurement of \(y_i\) is drawn from a Gaussian distribution with a mean \(y_0(x_i)\) and STD \(\sigma_i\)
  - Thus the probability of making the observed measurement of \(y_i\) (with \(\sigma_i\)) about the “true” mean value \(y_0(x_i)\) is given by
    \[
    P_i = \frac{1}{\sigma_i\sqrt{2\pi}} \exp\left\{-\frac{1}{2} \left[\frac{y_i - y_0(x_i)}{\sigma_i}\right]^2\right\}
    \]
• The probability of making a set of N measurements of \( y_i \)

\[
P(A_o, B_o) = \pi P_i = \prod \left( \frac{1}{\sigma_i \sqrt{2\pi}} \right) \exp \left\{ -\frac{1}{2} \sum \left[ \frac{y_i - y_o(x_i)}{\sigma_i} \right]^2 \right\}
\]

where the \( \sigma_i^2 \) are weighting factors

• Likewise for any estimated values of A and B the probability of obtaining the N values of \( Y_i \) is

\[
P(A, B) = \prod \left( \frac{1}{\sigma_i \sqrt{2\pi}} \right) \exp \left\{ -\frac{1}{2} \sum \left[ \frac{y_i - y(x_i)}{\sigma_i} \right]^2 \right\}
\]

• Fundamental assumption: the observed set of \( (x_i, y_i) \) is more likely to come from the parent distribution \( y_0(x) = A_0 + B_0x \) than any other with different coefficient and \( P(A_0, B_0) \) is the maximum probability one can have; consequently the most likely estimates of A and B are those that maximize \( P(A, B) \)
Physics 433 – Lecture 7

• Maximizing P(A,B) reduces to minimizing Chi squared

\[ \chi^2 = \sum_{i=1}^{N} \left[ \frac{1}{\sigma_i^2} (y_i - A - Bx_i) \right]^2 \]

• We now find values of A and B that minimize the weighted sum of the squares of the deviations of \( \chi^2 \) by taking the derivatives of \( \chi^2 \) with respect to A and B and setting them equal to zero

\[ \frac{\partial}{\partial A} \chi^2 = \frac{\partial}{\partial A} \sum \left[ \frac{1}{\sigma_i^2} (y_i - A - Bx_i)^2 \right] = -2 \sum \left[ \frac{1}{\sigma_i^2} (y_i - A - Bx_i) \right] = 0 \]

\[ \frac{\partial}{\partial B} \chi^2 = \frac{\partial}{\partial B} \sum \left[ \frac{1}{\sigma_i^2} (y_i - A - Bx_i)^2 \right] = -2 \sum \left[ \frac{x_i}{\sigma_i^2} (y_i - A - Bx_i) \right] = 0 \]
• The solutions, written in determinant form are

\[
A = \frac{1}{\Delta} \left| \begin{array}{cc}
\frac{\sum y_i}{\sigma_i^2} & \frac{\sum x_i}{\sigma_i^2} \\
\frac{\sum x_i y_i}{\sigma_i^2} & \frac{\sum x_i^2}{\sigma_i^2}
\end{array} \right| = \frac{1}{\Delta} \left( \sum \frac{x_i^2}{\sigma_i^2} \sum \frac{y_i}{\sigma_i^2} - \sum \frac{x_i}{\sigma_i^2} \sum \frac{x_i y_i}{\sigma_i^2} \right)
\]

\[
B = \frac{1}{\Delta} \left| \begin{array}{cc}
1 & \frac{\sum y_i}{\sigma_i^2} \\
\frac{\sum x_i y_i}{\sigma_i^2} & \frac{\sum x_i y_i}{\sigma_i^2}
\end{array} \right| = \frac{1}{\Delta} \left( \sum \frac{1}{\sigma_i^2} \sum \frac{x_i y_i}{\sigma_i^2} - \sum \frac{x_i}{\sigma_i^2} \sum \frac{y_i}{\sigma_i^2} \right)
\]

\[
\Delta = \left| \begin{array}{cc}
\frac{1}{\sigma_i^2} & \frac{\sum x_i}{\sigma_i^2} \\
\frac{\sum x_i}{\sigma_i^2} & \frac{\sum x_i^2}{\sigma_i^2}
\end{array} \right| = \sum \frac{1}{\sigma_i^2} \sum \frac{x_i^2}{\sigma_i^2} - \left( \sum \frac{x_i}{\sigma_i^2} \right)^2
\]
Least squares fit to a straight line \( y = A + Bx \) we have seen results from minimizing chi squared,

\[
\chi^2 = \sum_1^N \left( \frac{1}{\sigma_i} (y_i - A - Bx_i) \right)^2
\]

which yeilds A and B in terms of the variables \( (x_i, y_i) \).

The uncertainties in the parameters A and B follows from the standard variance calculations

\[
\frac{\partial A}{\partial y_j} = \frac{1}{\Delta} \left( \frac{1}{\sigma_j} \sum x^2 - \frac{x^2}{\sigma^2} \sum x^2 \right) \quad \frac{\partial B}{\partial y_j} = \frac{1}{\Delta} \left( \frac{x_j}{\sigma^2} \sum \frac{1}{\sigma^2} - \frac{1}{\sigma^2} \sum x_i \right)
\]

Yeilding

\[
\sigma^2_A \approx \sum_{j=1}^N \sigma_j^2 \left[ \frac{1}{\sigma_j^4} \left( \sum \frac{x^2}{\sigma^2} \right) - \frac{2x_j}{\sigma_j^4} \sum \frac{x^2}{\sigma^2} \sum \frac{x_i}{\sigma_i^2} + \frac{x^2}{\sigma_j^2} \left( \sum \frac{x_i}{\sigma_i^2} \right)^2 \right] = \frac{1}{\Delta} \sum \frac{x^2}{\sigma_i^2}
\]

\[
\sigma^2_B \approx \sum_{j=1}^N \sigma_j^2 \left[ \frac{x^2_j}{\sigma_j^4} \left( \sum \frac{1}{\sigma_i^2} \right)^2 - \frac{2x_j}{\sigma_j^4} \sum \frac{1}{\sigma_i^2} \sum \frac{1}{\sigma_i^2} - \frac{1}{\sigma_j^2} \left( \sum \frac{x_i}{\sigma_i^2} \right)^2 \right] = \frac{1}{\Delta} \sum \frac{1}{\sigma_i^2}
\]
Phys 433 - Lecture 7

- Propagation of errors in case of dependent variables that are functions of one or more measured variable or combining standard deviations of uncertainties to estimate a resultant uncertainty
- Consider the function $x = f(u, v, ...) \text{ and its most probable value } \bar{x} = f(\bar{u}, \bar{v}, ...)$
  - Individual measurements yield $x_i = f(u_i, v_i, ...)$
  - In limit of infinite measurements the variance is given by
    $$\sigma_x^2 = \lim_{N \to \infty} \frac{1}{N} \sum (x_i - \bar{x}_i)^2$$
  - The deviations of $x_i$ - $x$ are given by
    $$x_i - \bar{x} \cong (u_i - \bar{u}) \left( \frac{\partial x}{\partial u} \right) + (v_i - \bar{v}) \left( \frac{\partial x}{\partial v} \right) + ...$$
- Consequently
  $$\sigma_x^2 \cong \lim_{N \to \infty} \frac{1}{N} \sum \left[ (u_i - \bar{u}) \left( \frac{\partial x}{\partial u} \right) + (v_i - \bar{v}) \left( \frac{\partial x}{\partial v} \right) + ... \right]^2$$
  $$\cong \lim_{N \to \infty} \frac{1}{N} \sum \left[ (u_i - \bar{u})^2 \left( \frac{\partial x}{\partial u} \right)^2 + (v_i - \bar{v})^2 \left( \frac{\partial x}{\partial v} \right)^2 + 2(u_i - \bar{u})(v_i - \bar{v}) \left( \frac{\partial x}{\partial u} \right) \left( \frac{\partial x}{\partial v} \right) + ... \right]$$

May 11, 2006  Henry Lubatti
The first two terms are simply the variances given by

\[ \sigma_u^2 = \lim_{N \to \infty} \left( \frac{1}{N} \sum (u_i - \bar{u}_i)^2 \right) \quad \sigma_v^2 = \lim_{N \to \infty} \left( \frac{1}{N} \sum (v_i - \bar{v}_i)^2 \right) \]

The third term is the covariance \( \sigma_{uv} \equiv \lim_{N \to \infty} \left( \frac{1}{N} \sum [(u_i - \bar{u})(v_i - \bar{v})] \right) \)

Using these definitions we have the error propagation eq.

\[ \sigma_x^2 \cong \sigma_u^2 \left( \frac{\partial x}{\partial u} \right)^2 + \sigma_v^2 \left( \frac{\partial x}{\partial v} \right)^2 + ... + 2\sigma_{uv} \left( \frac{\partial x}{\partial u} \right) \left( \frac{\partial x}{\partial v} \right) + ... \]

The first two terms are the averages of the squares of the deviations in \( x \) produced by uncertainties in \( u \) and \( v \)

The cross term, covariance, vanishes if the variables are uncorrelated.

May 11, 2006

Henry Lubatti
• For the special case where all sigma are equal $\sigma = \sigma_i$ we have

\[
A = \frac{1}{\Delta'} \left| \begin{array}{cc} \Sigma y_i & \Sigma x_i \\ \Sigma x_i y_i & \Sigma x_i^2 \end{array} \right| = \frac{1}{\Delta'} \left( \Sigma x_i^2 \Sigma y_i - \Sigma x_i \Sigma x_i y_i \right)
\]

\[
B = \frac{1}{\Delta'} \left| \begin{array}{cc} N & \Sigma y_i \\ \Sigma x_i & \Sigma x_i y_i \end{array} \right| = \frac{1}{\Delta'} \left( N \Sigma x_i y_i - \Sigma x_i \Sigma y_i \right)
\]

\[
\Delta' = \left| \begin{array}{cc} N & \Sigma x_i \\ \Sigma x_i & \Sigma x_i^2 \end{array} \right| = N \Sigma x_i^2 - (\Sigma x_i)^2
\]

• Chi squared, $\chi^2$, can be used to test the "goodness of fit"
Chi squared test for goodness of fit

- In the least squares fitting procedure we minimize the weighted sum of the squares of deviations of the data points, \( y_i \), from the fitting function, \( y(x_i) \)
- The variance of the fit \( s^2 \) is an estimate of the data, \( \sigma^2 \)

The variance of the fit is given by

\[
s^2 = \frac{1}{d} \sum w_i [y_i - y(x_i)]^2
\]

where \( d = N - m \) is the number of degrees of freedom, \( N \) the total number of data points, \( m \) the number of parameters in the fit and \( w_i = \frac{1/\sigma_i^2}{\left(\frac{1}{N}\right) \sum \left(\frac{1}{\sigma_i^2}\right)} \) the weights for the \( i^{th} \) data point.

- The variance is related to Chi Squared
- In our fits \( y(x_i) = A + Bx_i \)
- From the above definitions we find the reduced

\[
\chi^2_d = \frac{\chi^2}{d} = \frac{s^2}{\langle \sigma_i^2 \rangle}
\]
The next step is to determine some estimate of the “goodness” of your fit

There are two tests often used: the linear correlation coefficient, $r$, and the Chi squared probability

- Linear Correlation Coefficient $r$, is a measure of how well the set of data points $(x_i, y_i)$ correlate
- For example, if you measure the force required to stretch a spring a distance $x$ and want to determine if force has the following dependence on distance, $F = A + Bx$ (in this case $B$ represents the spring constant $k$).
- In this example you would find a linear correlation between the variables; however if you made a study of the force required to stretch a spring a given distance as a function of time, aside from the statistical fluctuations in your measurements there would be no reproducible correlations between the variables
Physics 433 – Lecture 7

- If the variables \((x_i, y_i)\) with \(y\) the dependent variable are described by \(y = A + Bx\) we examine what we obtain when we treat \(x\) as the dependent variable, that is \(x = a + by\)
  
  - \(A\) and \(B\) are given by
    
    \[
    A = \frac{1}{\Delta'} \begin{vmatrix} \Sigma y_i & \Sigma x_i \\ \Sigma x_i, y_i & \Sigma x_i^2 \end{vmatrix} - \frac{1}{\Delta} (\Sigma x_i^2 \Sigma y_i - \Sigma x_i \Sigma x_i y_i)
    \]
    
    \[
    B = \frac{1}{\Delta'} \begin{vmatrix} N & \Sigma y_i \\ \Sigma x_i & \Sigma x_i \Sigma y_i \end{vmatrix} = \frac{1}{\Delta} (N \Sigma x_i y_i - \Sigma x_i \Sigma y_i)
    \]
    
    \[
    \Delta' = \begin{vmatrix} N & \Sigma x_i \\ \Sigma x_i & \Sigma x_i^2 \end{vmatrix} = N \Sigma x_i^2 - (\Sigma x_i)^2
    \]

  - The coefficients \(a(x_i, y_i)\) and \(b(x_i, y_i)\) are obtained by interchanging \(x\) and \(y\) in the equations above

  - If there is complete correlation between \(x\) and \(y\) we can determine a relationship between the coefficients that we can obtain by writing \(y = -a/b + x/b = A + Bx\); equating the coefficients gives
    
    \[A = -a/b\] and \(B = 1/b\) so for complete correlation, \(bB = 1\) and no correlation is when \(bB = 0\)

- We define the linear correlation coefficient \(r = \sqrt{bB}\)
  
  - Substituting \(b(x_i, y_i)\) and \(B(x_i, y_i)\)
  
  where \((0 < r < 1)\)