

Matt's Error Analysis Slides

7/7/09 Group Meeting

Data Reproducibility and Error Bars

- Minimum of 3 data points to report an error bar
- Error bars should represent a confidence interval
- Individual data points within a data set should have different error bars

$$s = \sqrt{\frac{\sum_{j=1}^N (X_j - \bar{X})^2}{N}} = \sqrt{\frac{\sum (X - \bar{X})^2}{N}} = \sqrt{\frac{\sum x^2}{N}} = \sqrt{(X - \bar{X})^2}$$

- $\pm S \rightarrow 68\%$ confidence interval, $\pm 1.96 * S \rightarrow 95\%$ confidence interval, etc.
- Normalized confidence interval

Significance of Data

- P-Value
 - Consider a data set in which you are repeating each data point 3 times. What is the probability that variation between the means of each data point would be the same if you repeated infinite times?
 - (ex) P-Value of 0.05 → corresponds to a 5% chance that you observed a difference between the mean data points even though the means would be the same if you took infinite samples
- F-Test
- Example in Minitab

Data Regression

- Method of Least Squares
 - Minimize the sum of squared residuals between your data points and you model prediction

$$S = \sum_{i=1}^n r_i^2$$

- Can be done in MS excel, however a full error report sometimes requires more sophisticated software
- Example in MS Excel