Week 2

Note Title 1/19/2007

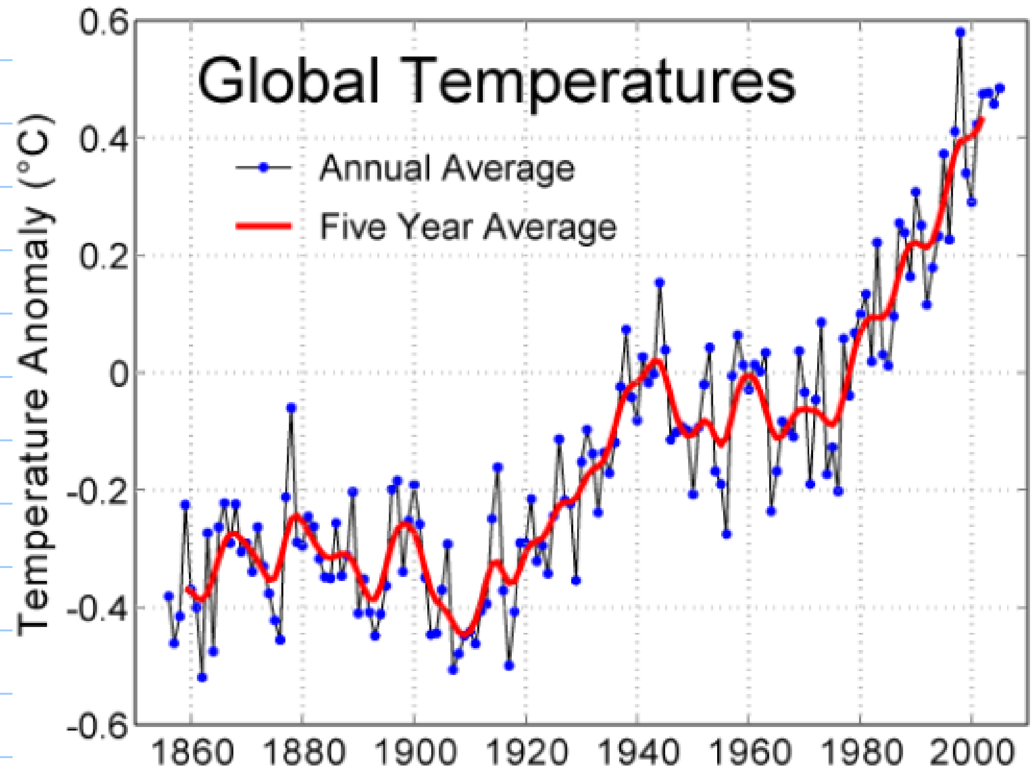
"Figures often beguile me, particularly when I have the arranging of them myself: in which case the remark attributed to Disraeli would often apply with justice and force: 'There are three kinds of lies: lies, damned lies and statistics'"-Mark Twain

Data reliability

- accuracy – deviation from accepted value
- precision – averaging improves
- error – deviation from true value

Types of error

- systematic errors (determinate errors)
- random errors (indeterminate errors)



Temp Increase

1.1 ± .4° F

in 20th

century

Assessing Accuracy

$$\text{Percent Error} = \frac{\text{measured value} - \text{true value}}{\text{true value}} \times 100\%$$

- If we know the true value we can determine a percent error
- If we don't know the true value we try to assess the accuracy of a group of measurements

Simple Average

$$\bar{x} = \frac{x_1 + x_2 + x_3 + x_4 + x_n}{N} = \frac{1}{N}(x_1 + x_2 + x_3 + x_4 + x_n) = \frac{1}{N}\sum_{i=1}^{N} x_i \qquad (2)$$

- in the absence of systematic error $\bar{x}$ approaches the true value
- if there is a systematic error no amount of measurements will give the true value

Uncertainty

$$\bar{x} = x_{best} \pm \delta x$$

Question?
What is the meaning of δx and how is it determined

Square-Root rule for counting experiments

$$\left(\begin{array}{c}\text{Averge number of events}\\ \text{in time } T\end{array}\right) = \nu \pm \sqrt{\nu}$$

Square Root Rule

10 events in time T $\qquad 10 \pm \sqrt{10}$

20 events in time $2T$ $\qquad 20 \pm \sqrt{20}$

$$\Rightarrow \text{ in time } T = \frac{20 \pm \sqrt{20}}{2}$$

$$= 10 \pm \sqrt{5}$$

$$\frac{\delta n_{2T}}{\delta n_T} = \sqrt{\frac{10}{5}} = \underline{\underline{\sqrt{2}}} \text{ improvement!}$$

Propagation of Uncertainties

Sums + differences

- Consider two quantities

$$x = x_{best} \pm \delta x$$

$$y = y_{best} \pm \delta y$$

Highest Value

$$x_{best} + y_{best} + \delta x + \delta y$$

Lowest value

$$x_{best} + y_{best} - (\delta x + \delta y)$$

is

$$q = x + y$$

$$= x_{best} + y_{best}$$

where

$$\delta q \approx \delta x + \delta y$$

Independent Uncertainties

- Consider the uncertainty in the sum

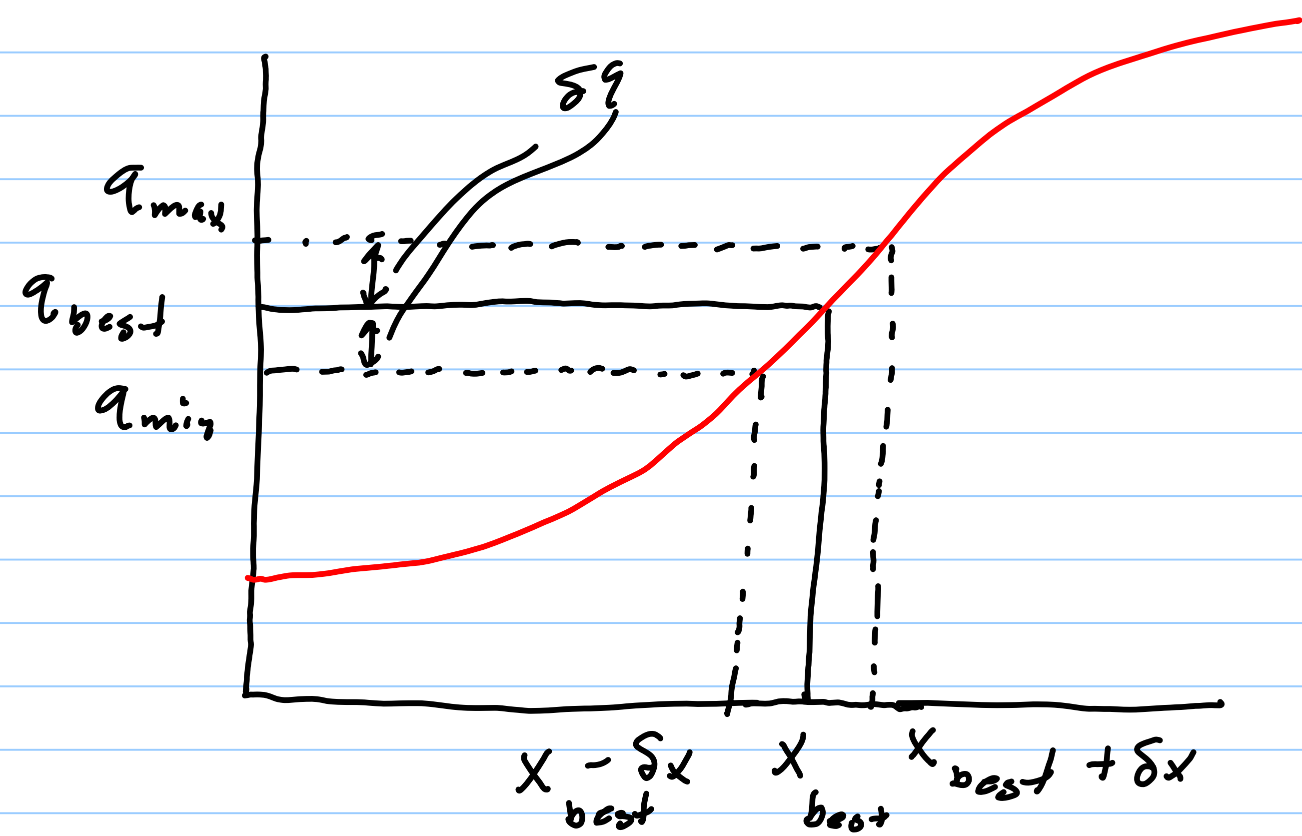
$$q = x + y \qquad \delta q = \delta x + \delta y$$

δq is an over estimate

We will see later that for a normal distribution

$$\delta q = \sqrt{\delta x^2 + \delta y^2}$$

Consider Arbitrary Functions of One Variable



$$\delta q = \left| \frac{dq}{dx} \right| \delta x$$

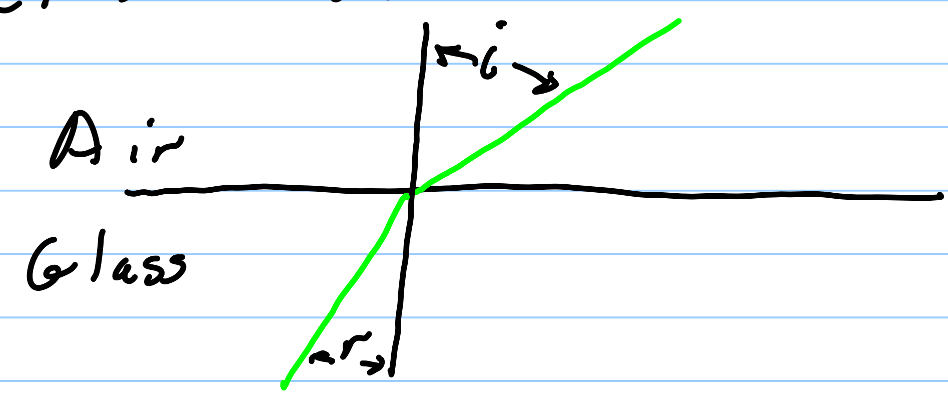
Suppose

$$q = x^n$$

$$\delta q = \left| \frac{dq}{dx} \right| \delta x = \left| n x^{n-1} \right| \delta x$$

$$\Rightarrow \frac{\delta q}{q} = |n| \frac{\delta x}{|x|}$$

Snell's Law



$$n = \frac{\sin i}{\sin r}$$

$$\Rightarrow \quad \frac{\delta n}{n} = \sqrt{\left(\frac{\delta \sin(i)}{\sin i}\right)^2 + \left(\frac{\delta \sin r}{\sin r}\right)^2}$$

General Formula for Error Propagation

$$q = q(x, \ldots z) \quad \text{any fcn of } x \ldots z$$

$$\delta q = \sqrt{\left(\frac{\partial q}{\partial x}\delta x\right)^2 + \cdots \left(\frac{\partial q}{\partial z}\delta z\right)^2} \quad \left(\begin{array}{c}\text{errors indep} \\ + \\ \text{random}\end{array}\right)$$

and

$$\delta q \leq \left|\frac{\partial q}{\partial x}\right|\delta x + \cdots \left|\frac{\partial q}{\partial z}\right|\delta z$$

Mean and Standard Deviation

Consider 5 measurements

$$71,\ 72,\ 72,\ 73,\ 71$$

$$x_{best} = \bar{x}$$

$$= \frac{71 + 72 + 72 + 73 + 71}{5}$$

$$= 71.8$$

$$\bar{x} = \frac{x_1 + x_2 + \cdots\cdots x_N}{N}$$

$$= \frac{\Sigma x_i}{N}$$

Define deviation

$$d_i = x_i - \bar{x}$$

Standard deviation

$$\sigma_x = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(d_i)^2} = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(x - \bar{x})^2}$$

Sample Standard deviation

$$\sigma_x = \sqrt{\frac{1}{N-1} \sum (x_i - \bar{x})^2}$$

note factor

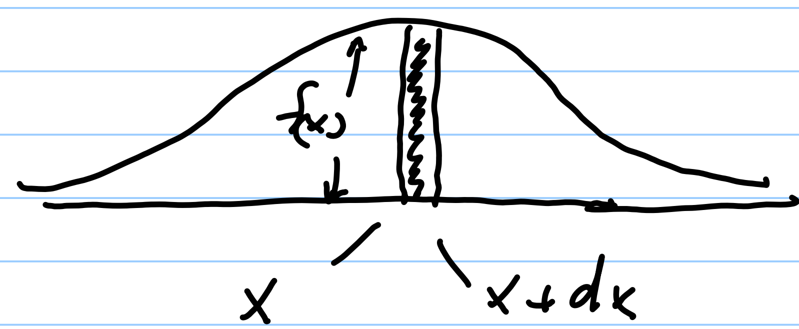
Central Limit Theorem

If independent variables have a finite variance the sum of these variables will show a normal distribution.

— Simulation —

Limiting Distribution $f(x)$

$f(x)\,dx$ = fraction of measurements between x and $x + dx$



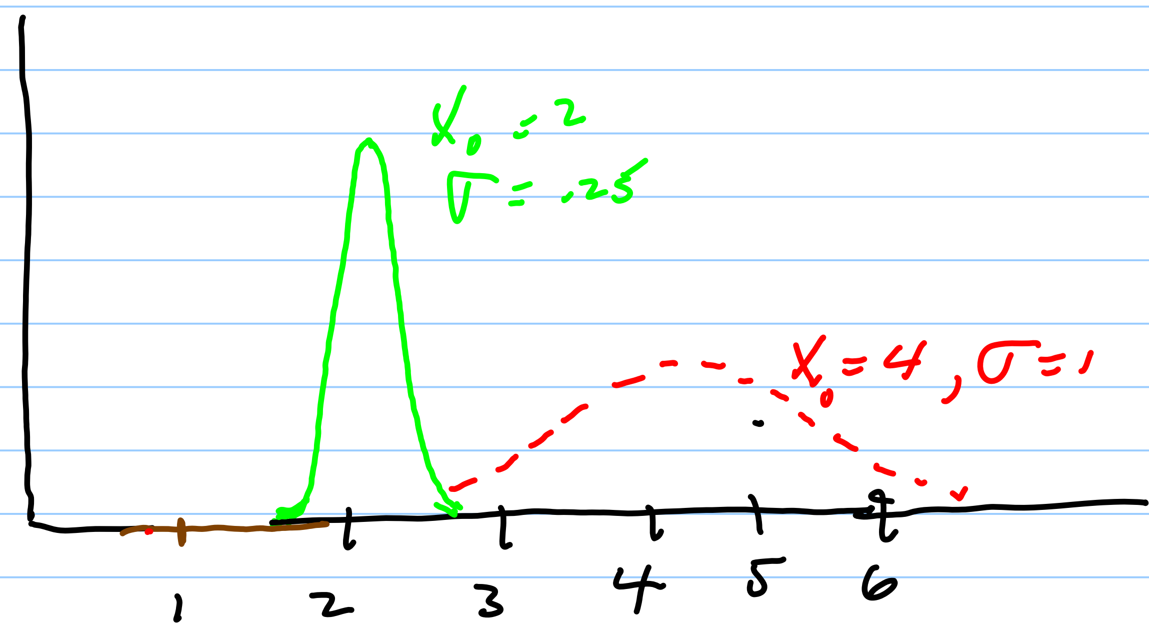
$$\int_{-\infty}^{\infty} f(x)\, d(x) = 1$$

Mean

$$\bar{x} = \int_{-\infty}^{\infty} x\, f(x)\, dx$$

Gauss or Normal Distribution

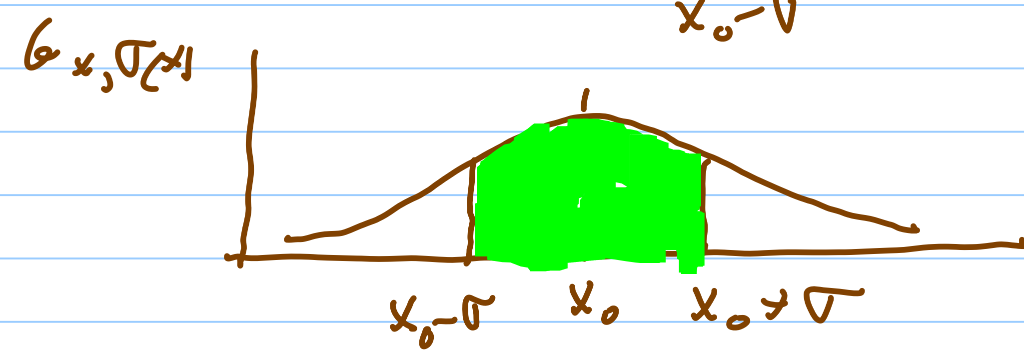
$$G_{x,\sigma}(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-(x-x_0)^2/2\sigma^2}$$



Standard Deviation 68% Confidence Limit

$$\int_a^b f(x)\,dx$$

$$\text{Prob}(\text{within } \sigma) = \int_{x_0-\sigma}^{x_0+\sigma} G_{x_0,\sigma}(x)\,dx$$



Reduces to

$$\text{Prob}(\text{within } \sigma) = \frac{1}{\sqrt{2\pi}} \int_{-1}^{1} e^{-z^2/2}\, dz$$

$$= .68$$

Acceptability of a measured answer

$$(\text{value of } x) = x_{best} \pm \sigma$$

$$t = \frac{x_{best} - x_{exp}}{\sigma}$$