Measurements, Predictions, Comparisons and Errors

Arne Freyberger
Operations Dept.
Accelerator Division
JLAB

April 17, 2012

Outline

1. Measurements and Their Uncertainties
2. Probability Distribution and Measurement Uncertainty
3. Why is everything Gaussian?: The Central Limit Theorem
4. Measured versus True, Systematic Uncertainties, Combining Uncertainties
5. Whoops, not everything is Gaussian: Correlated Quantities
6. Particle/Nuclear Physics Approach to Uncertainties and some Accelerator Based Examples
   - Emittance and TWISS parameter measurement: Quad Scan and Wire Scanners
   - Using Monte Carlo to Estimate SRF Cryogenic Loads
   - Maximum Likelihood: For when all else fails
7. References and Summary

Measurement Overload

An accelerator is an overwhelming source of measurements:

Beam	position, size, energy, intensity, bunch length, energy spread, polarization
RF systems	gradient, frequency, phase, cathode current
Cryogenics	Pressure, flow, temperature, valve location
Magnets	Current, Voltage, Temperature, LCW flow
Vacuum	Pressure, particle species
...	...

Understanding the uncertainty (or error) in these measurements is vital in the correct interpretation of the measurement/system.

Measurement Uncertainty

Every measurement has uncertainty. This is sometimes referred to as the *instrumental uncertainty* and its value is inherent to the device, This defines the precision of the device.

Bit Resolution: 8bit ADC, 8bits = 256, 5V/256bits = 0.02V/bit

Circuit Noise: Electronic noise sources ($\sqrt{4kT \cdot R \cdot B}$ thermal noise, shot noise, flicker noise) that limit the precision of the measurement.

Scale Resolution: Old school this meant the graduations on the scale. In the digital age, this refers to the significant digits promoted to the User Interface or stored in the archiver. Often these are truncated to reserve space at a loss of precision.

... ...

Uncertainties

In general the uncertainty estimate should strive to equate to the root mean square (RMS)† deviation of an infinite set of measurements. A measurement (the mean) and its uncertainty (RMS) is usually denoted as:

$$x \pm \Delta x$$

For a set of measurements $(x_1, x_2, \ldots, n)$,

$$x = \bar{x} = \frac{1}{n} \sum x_i$$

$$\Delta x = RMS = \sqrt{\frac{1}{n-1} \sum (x_i - \bar{x})^2}$$

Note that these equations hold for all types of distributions.

† Throughout this talk RMS refers to the root mean standard deviation.

Probability Distributions

If the uncertainty of a measurement is defined as the RMS of a set of measurements (or its equivalent), then the uncertainty for a probability distribution is straight forward to determine via the analog continuous definitions:

$$\mu = \int_{-\infty}^{\infty} xP(x)dx$$

$$\sigma^2 = \int_{-\infty}^{\infty} (x-\mu)^2 P(x)dx = \int_{-\infty}^{\infty} x^2 P(x)dx - \mu^2$$

For the record:

$$1 = \int_{-\infty}^{\infty} P(x)dx$$

The Gaussian Distribution Function

The most useful or common distribution is the Gaussian probability distribution:

$$P(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-0.5(\frac{x-\mu}{\sigma})^2}$$

where, $\bar{x} = \mu$ and $RMS = \sigma$.

Using the Gaussian distribution as a guide, an alternative definition of the uncertainty can be made:

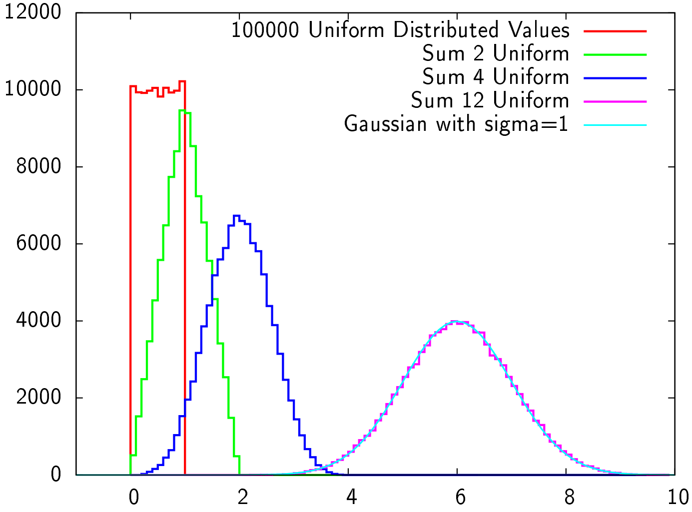
Δx defines a region such that the *true* value, $\mathcal{X}$ has the probability of residing in the range:

$$(\bar{x} - \Delta x)) < \mathcal{X} < (\bar{x} + \Delta x) \text{ of } 68.3\%$$

With the most probable value for $\mathcal{X}$ being $\bar{x}$.

Central Limit Theorem

The sum of **independent** random numbers (**of any distribution**) becomes Gaussian distributed as $N \rightarrow \infty$



Everything is Gaussian

Integrating the Gaussian distribution over a limited range is best done numerically. Luckily these days this is a built in function in most spreadsheets and numerical libraries. The function is called the *error function* and is defined as:

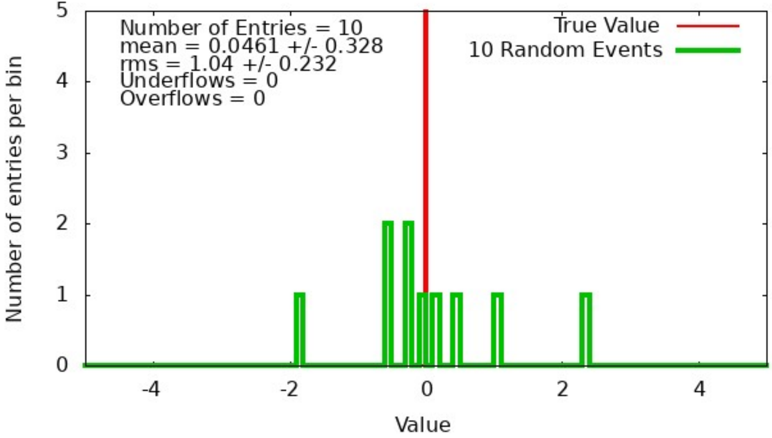
$$\mathrm{erf}(x) = \frac{2}{\pi}\int_0^x e^{-t^2}\,dt$$

$\mathrm{erf}(\frac{a}{\sigma\sqrt{2}})$ is the probability that a measurement lies between $-a$ and a.

nσ	Area (%)
1	68.27
1.645	90.00
1.960	95.00
2	95.45
2.576	99.00
3	99.73
3.290	99.90
4	99.99
5	100.00

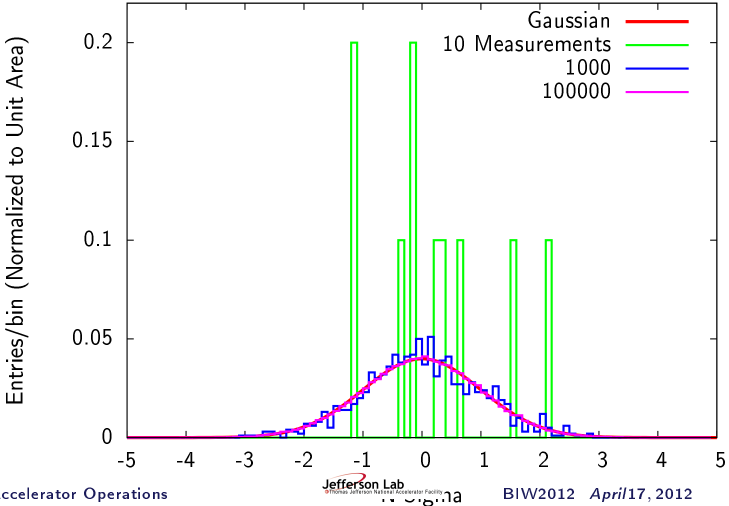
True, Expected or Modeled Value

- The resolution of the measurement, RMS or σ, is fixed by the experimental equipment. (and not improved by statistics)
- The measured value and its uncertainty provides a **estimate** for the range of the *true* value. We never **know** what the true value is.



What does statistics buy you?

Repeating a measurement N times will not improve the experimental uncertainty, but it will improve our estimate of the uncertainty and the measured mean.



Mean and Width Uncertainty

For N measurements, the mean, $\bar{x}$, is determined with the following uncertainty:

$$\delta\bar{x} = \frac{\sigma}{\sqrt{N}}$$

And the width, σ or RMS, is determined with uncertainty:

$$\delta\sigma = \frac{\sigma}{\sqrt{2N}}$$

Remember the true value, $\mathcal{X}$, is most probably $\bar{x}$, so the more precise $\bar{x}$ is determined the better the determination of $\mathcal{X}$.

Example: Beam Position

At CEBAF the nominal Beam Position Monitor resolution is quoted at $50\mu m$. A set of experiments, parity scattering, requires that the beam position for positive spin aligned electrons be within nanometers of the negative spin aligned electrons.

How many measurements of beam position are required to achieve *nm* uncertainty on the average beam position?

$$\delta\bar{x} = 10^{-9}m = \frac{50 \times 10^{-6}m}{\sqrt{N}}$$

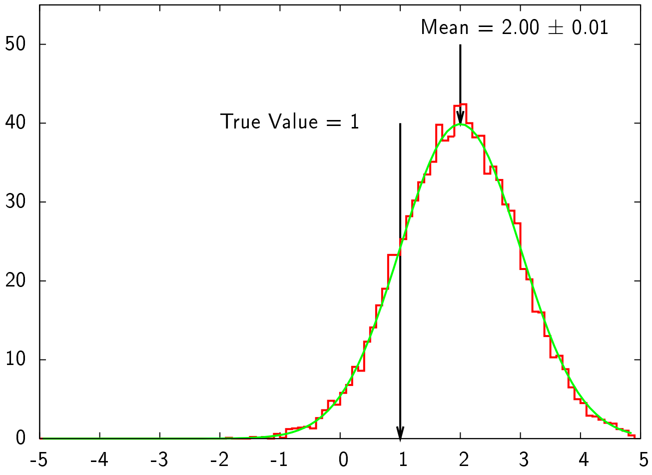
$$N = 2.5 \times 10^{9}$$

If measurements are made at 960Hz, how long will it take the to achieve this goal?

$$T(days) = \frac{2.5 \times 10^{9}}{960}\frac{1}{3600 \times 24} = 30.1\mathrm{days}$$

Systematic Error vs. Uncertainty

If through some independent method it is determined that the reported measured mean is **not** the most probable value for the true value, $\mathcal{X}$. Then the measurement is said to have a systematic error.



Combining Uncertainty Terms
Uncorrelated Terms

If the measurement is a function of several independent parameters:

$$\bar{x} = \bar{x}(\alpha_1, \alpha_2, \ldots, \alpha_n)$$

The total uncertainty is the *sum of the squares* of the variation (partial derivative) with respect to each independent parameter.

$$\Delta\bar{x} = \sqrt{\sum_i \left(\frac{\partial \bar{x}}{\partial \alpha_i} \Delta\alpha_i\right)^2}$$

Often working with relative uncertainty is more straight forward:

$$\frac{(\Delta\bar{x})^2}{\bar{x}^2} = \frac{1}{\bar{x}^2} \sum_i \left(\frac{\partial \bar{x}}{\partial \alpha_i} \Delta\alpha_i\right)^2$$

Example: Combining Uncorrelated Uncertainties

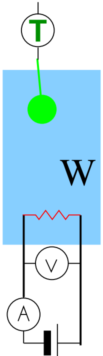
Determining Heat Capacity of W slug

Using thermometry and a heater, what is the expected error on determining the Heat Capacity of the W slug?

$$C_m(Joules/^\circ K) = I \cdot V \cdot \delta t/\Delta T$$

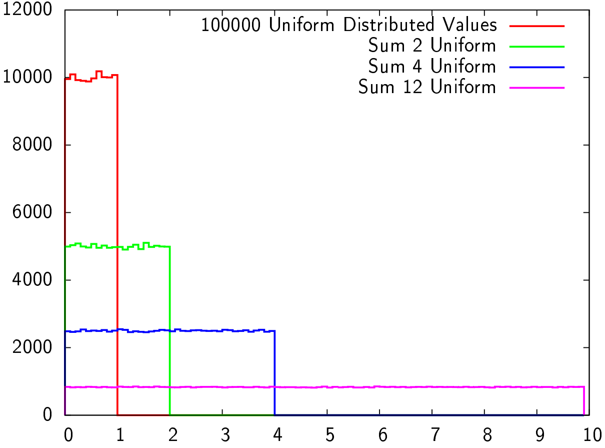
$$\frac{\Delta C_m}{C_m} = \sqrt{(\frac{\Delta I}{I})^2 + (\frac{\Delta V}{V})^2 + (\frac{\Delta(\Delta T)}{\Delta T})^2}$$

Quantity	Expected Value	Uncertainty	Relative Uncertainty (%)
ΔT ($^\circ$K)	10	0.025	0.25
I_Ω (A)	15	0.0135	0.1
V_Ω (V)	75	0.006	0.01
Total			0.27



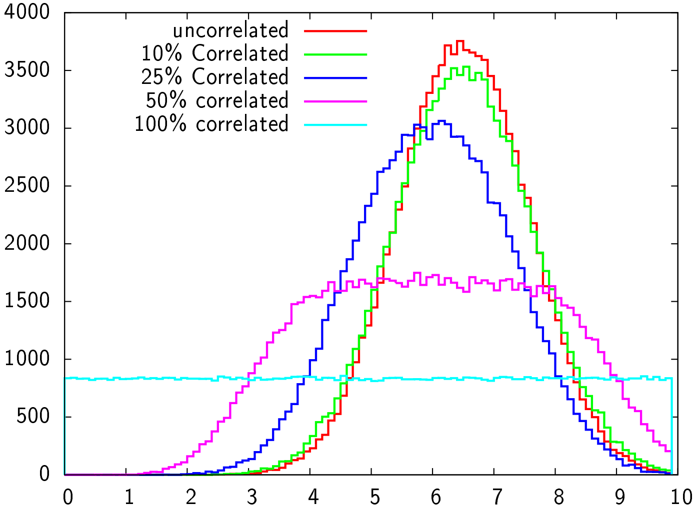
Correlated Uncertainties

What happens to the Central Limit Theorem in the presence of correlation between variables? 100% Correlation between random sets:



How much Correlation does it take?

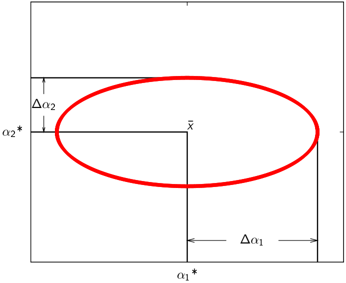
Combine 12 sets of 100,000 measurements with a fraction of the 100,000 sorted before being added to the total.



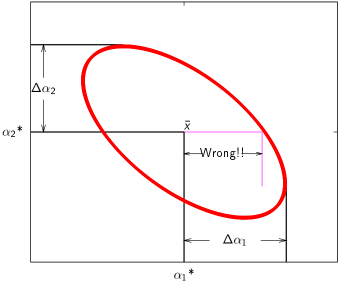
Graphical Interpretation of Uncertainty

Plots represents the 68.3% contour, or 1σ, on the probability distribution. Measurement if $\bar{x}$ depends of two variables, α_1 and α_2.

Uncorrelated



Correlated



$$\sigma_x^2 \simeq \sigma_u^2\left(\frac{\partial x}{\partial u}\right)^2 + \sigma_v^2\left(\frac{\partial x}{\partial v}\right)^2 + 2\sigma_{uv}^2\frac{\partial x}{\partial u}\frac{\partial x}{\partial v} + \cdots$$

The Particle/Nuclear physicists Approach to Uncertainties/Data Analysis

- Assume everything is Gaussian
- Uncertainties go as $\sqrt{N}$
- Variables are independent.
  - Many statistics/data analysis software libraries and tools are available.
- Non-Gaussian highly correlated situations are handled on a case by case basis starting from first principles.
  - Uncertainties can be estimated by developing problem specific Monte Carlos.
  - Maximum Likelihood Method

TWISS Parameter measurement: Quad Scan

This is common measurement on Accelerators. The process is as follows:

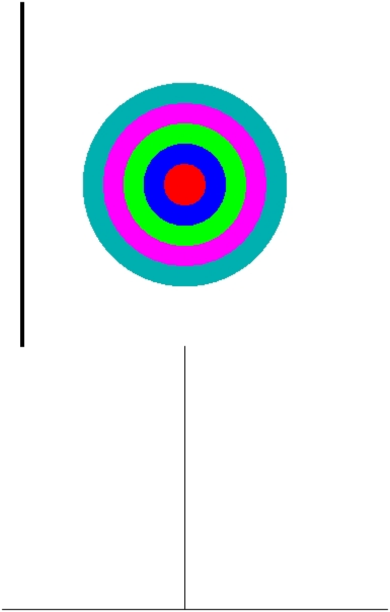
1. Measure the transverse beam profile, wire scanner (CEBAF) or viewer
2. Change optics
3. repeat steps 1 and 2

Sources of Uncertainty

- Uncertainty in the SEM measurement or PMT count.
- Uncertainty in transverse width
- Uncertainty in the transport matrix from Quad to Wire Scanner (Viewer)

  Magnetic field model, negligible?
  Element location, negligible?

Wire Scanner



Signal Source can be either detecting the SEM off the wire or by detecting the scattered particles with a photomultiplier.

SEM

- Signal level small
- Noise level high
- Width resolution of $\mathcal{O}(10\mu m)$

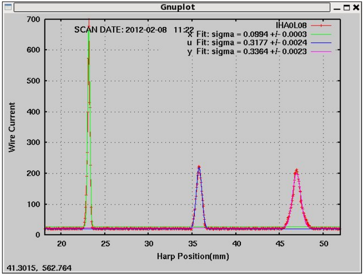
PMT

- If PMT in counted mode, $\sqrt{N}$ statistics on each measurement
- Properly configure width resolution can be as low as $\mathcal{O}(1\mu m)$

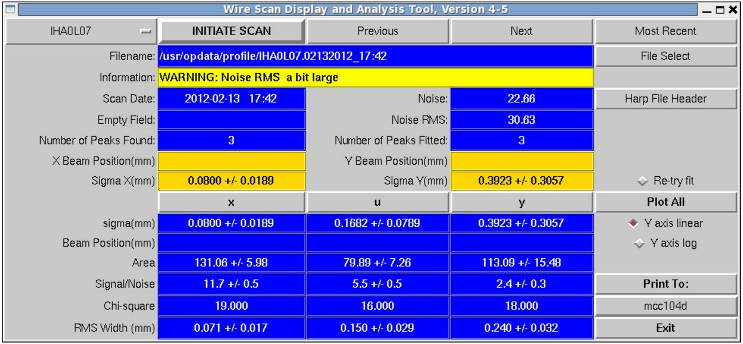
Secondary Emission Wire Scanner: Typical

1. Signal Noise is determined for each harp scan by averaging the signal term at the beginning and end of the data file.
2. Peaks are located via pattern recognition software
3. RMS deviations are determined (along with their uncertainties) and are used as starting points for a fit to the data.
4. Standard Least Square Fit assuming a Gaussian shape is performed.
5. Fitted widths and their uncertainties are extracted.

- $\text{RMS}_X = 103 \pm 20\mu\text{m}$
- $\sigma_\text{X} = 99.4 \pm 0.3\mu\text{m}$
- $\text{RMS}_X = 362 \pm 36\mu\text{m}$
- $\sigma_\text{Y} = 336.4 \pm 2.3\mu\text{m}$



Secondary Emission Wire Scanner: Noisey



- $\mathrm{RMS}_X = 71 \pm 17 \mu\mathrm{m}$
- $\sigma_\mathrm{X} = 80 \pm 19 \mu\mathrm{m}$
- $\mathrm{RMS}_X = 240 \pm 32 \mu\mathrm{m}$
- $\sigma_\mathrm{Y} = 390 \pm 300 \mu\mathrm{m}$

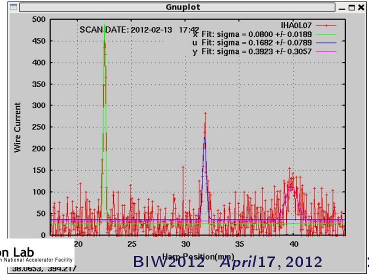
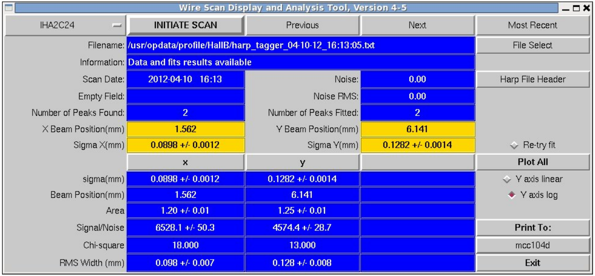
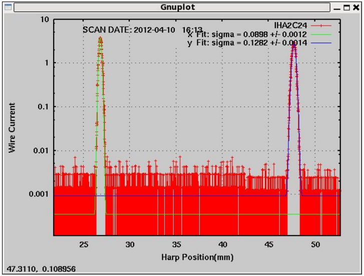


Photo-multiplier Wire Scanner: $\sqrt{N}$

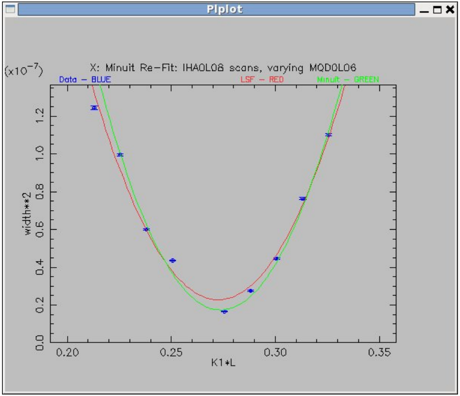


- $\mathrm{RMS}_X = 98 \pm 7\mu\mathrm{m}$
- $\sigma_\mathrm{X} = 89.8 \pm 1.2\mu\mathrm{m}$
- $\mathrm{RMS}_X = 128 \pm 8\mu\mathrm{m}$
- $\sigma_\mathrm{Y} = 128.2 \pm 1.4\mu\mathrm{m}$



Quad Scan Analysis

- A fit 2^{nd} order polynomial fit to $1/f$ versus σ^2 is performed.
    - First a straight forward LSF is performed to provide initial seeds to MINUIT.
    - The final results are derived by invoking MINUIT which provides a more complete exploration of the parameter space.
- The fitted parameters with uncertainties, A,B & C, are then used to extract ε, β and α and their uncertainties.

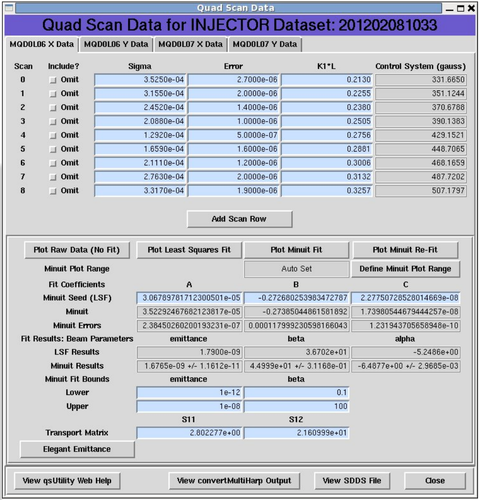


Quad Scan Example

- $\varepsilon = (1.677 \pm 0.012) \times 10^{-9}$ m-rad
- $\beta = 44.50 \pm 0.31$ m
- $\alpha = -6.488 \pm 0.0030$

Why Such a Busy Screen?

1. Wire Scanner Results suspect: Remove bad measurements or adjust width errors
2. Model suspect: Allow user to adjust Transport Matrix values
3. Report intermediate (A,B,C) and final results($\varepsilon, \ldots$).



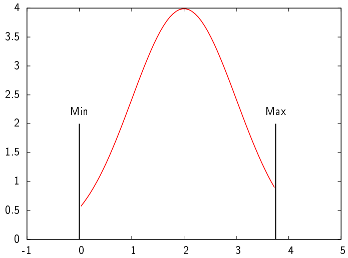
The Monte Carlo Method: Uncertainty Estimate for non-Gaussian distributions

- In cases where the $\frac{\partial f(x)}{\partial x}$ is challenging to calculate
- When the distributions are non-Gaussian, distributions with hard edges (min and max for example.)

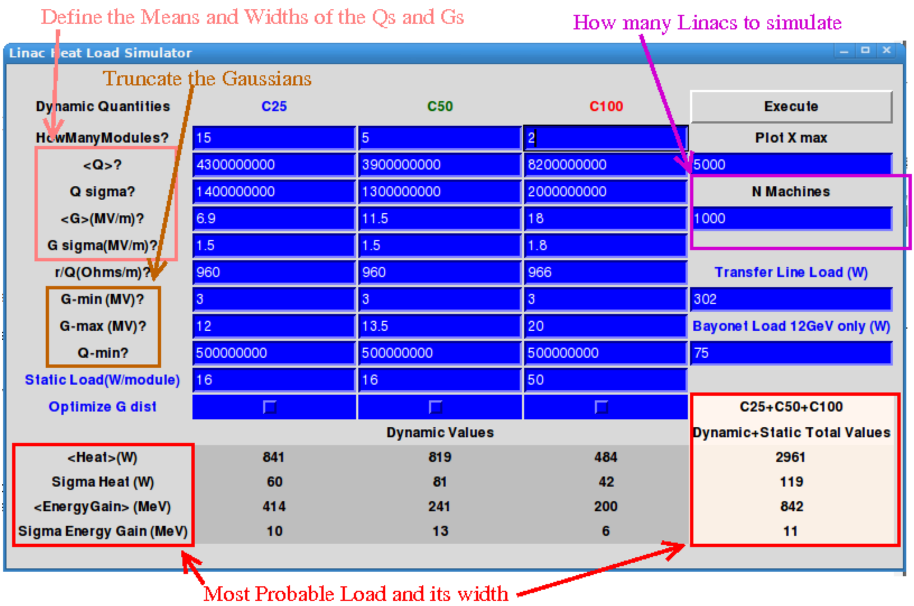
Given a random distribution of gradients, Qs, r/Q, with hardware limits on Gradient and Qs, what is the expected cryogenic heatload, $\mathcal{H}$ and its uncertainty.

$$\mathcal{H} \propto \frac{\mathcal{G}^2}{\mathrm{Cavity}_{R/Q}\mathcal{Q}}$$

where $\mathcal{G}$ and the cavity $\mathcal{Q}$ are distributions generated to reflect reality.



Linac Heat Load: Monte Carlo



Maximum Likelihood

See Orear CLNS82/511

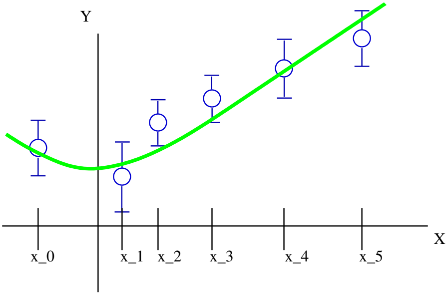
$\mathcal{L}(\alpha) = \prod_{i=1}^{N} f(\alpha, x_i)$ The likelihood function is the joint probability density of getting a particular experimental result, $x_1, \ldots, x_n$, assuming $f(\alpha, x)$ is a normalized distribution function: $\int f(\alpha, x)dx = 1$. Note, normalization is the only requirement on $f(\alpha, x)$.

$w \equiv \ln \mathcal{L}(\alpha_1, \ldots, \alpha_M)$ Define w to be the log of the likelihood function for M parameters to be determined.

$\frac{\partial w}{\partial \alpha_i}|_{\alpha_i=\alpha_i^*} = 0$ Yields values for α_i^* that maximize $\mathcal{L}$.

$\Delta\alpha = (-\frac{\partial^2 w}{\partial \alpha^2})^{-\frac{1}{2}}$ Maximum Likelihood Uncertainty

Numerical Maximum Likelihood

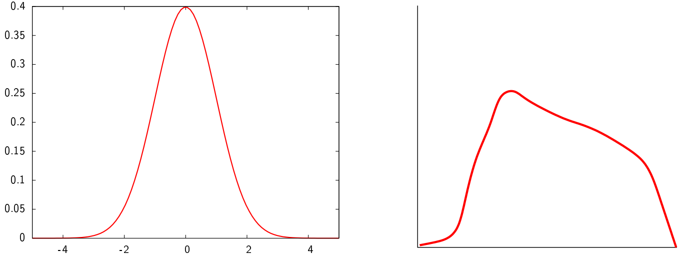


Let's take a set of measurements so that for each x_i we obtain a measured value and uncertainty, $y_i \pm \Delta y_i$. We have a hypothesis that the function $\bar{y}(x)$ represents the data.

Probability Distribution function: If Uncertainties are Gaussian

$$f(y_a, \bar{y}(x_a, \alpha_j, \ldots)) = \frac{1}{\sqrt{2\pi}\sigma_{y_a}} e^{-0.5 \frac{(y_a - \bar{y}(x_a))^2}{\sigma_{y_a}^2}}$$

Probability Density Function Shape can be arbitrary



In this case $\mathcal{L} = \prod f(y, \bar{y}(x, \alpha))$ is calculated numerically, in other words, represent each f as a normalized histogram and just multiple all the individual histograms together.

Note: that you will also have to step through the parameter space, α's to form the $\mathcal{L}(\alpha)$ distribution. We have computers!

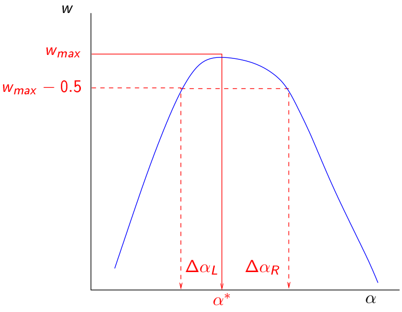
Numerical Maximum Likelihood

Non-Gaussian parameter and uncertainty

With $\mathcal{L}(\alpha)$ determined numerically, numerically take the log:

$$w(\alpha_i) = \ln \mathcal{L}(\alpha_i)$$

One dimensional Case:

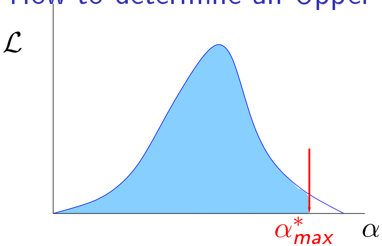


- α^* is the value of α for which w is a maximum.
- The region that represents 68.3% of the probability distribution is determined by the value of α the corresponds to: $w_{max} - 0.5$.
- The left and right uncertainties are often unequal.

A comment on Upper/Lower Limits

When the uncertainty is very large, sometimes it is best to quote that the most probable value is below (or above) some value at the XX% confidence level. The choice of a value for the limit is somewhat arbitrary, although I think these days most people quote 95% CL.

How to determine an Upper (or Lower) limit



- Pick the confidence level you wish to quote, i.e. 95%
- Find the limit of integration, *UL*, such that: $\int_{-\infty}^{UL} \mathcal{L}(\alpha)d\alpha = 0.95$

Value α^* has a 95% probability of being less than α^*_{max}.

References: In preferential order

Orear *Notes on Statistics for Physicists, Revised*, `CLNS 82/511` This is a concise, complete, concrete text that explains statistics and uncertainties all derived within a Maximum Likelihood framework. It contains very informative examples.

Hoffmann *Measurement, statistics, and errors* CAS2008 pg: 157-177. Concise treatment of noise terms.

James *MINUIT Manual*, Function Minimization and Error Analysis, FORTRAN and C++ implementations, with wrappers for JAVA, Python, Perl.

Bevington *Data Reduction and Error Analysis for The Physical Sciences*, McGraw-Hill Companies; 1992 ISBN-10: 0079112439, ISBN-13: 978-0079112439

Box *et al.* *Statistics for Experiments* Wiley-Interscience; 2 edition (May 31, 2005) ISBN-10: 0471718130 ISBN-13: 978-0471718130

Caria *Measurement Analysis* World Scientific Publishing Company (March 2001) ISBN-10: 1860942318 ISBN-13: 978-1860942310

Sandbox Summary

No one owns the truth!

Model calculations/simulations have uncertainties, just as well as physical measurements.

Much beam time has been spent trying to reconcile measurement & prediction only to find that it was the model that was wrong!

Three parameters define a measurement

1. Most probable value: α^*
2. The 68.3% uncertainty bounds: $\pm\Delta\alpha$
3. The unit

An honest comparison between model/theory and measurement cannot take place unless **all** uncertainties, measurement and model, are accounted for.

The Other Summary

Everything is Gaussian!! See Central Limit Theorem.

Life is a probability distribution, with a very high probability of being a Gaussian distribution.

- Uncertainty determination starts at the signal source.
- Care must be taken with regard to error propagation from the source to the final result in order to arrive with the correct value for the uncertainty.
  - It is easier to determine α^*, than it is to determine $\Delta\alpha$. (Which is why the uncertainty is often not presented).
- Standard software libraries and applications, old and new, are available to handle standard cases. Trust but verify!
- When in doubt or lazy invoke MINUIT to do the work for you.
- If necessary, resort to first principles and write down or numerically determine $\mathcal{L}$.