χ^2 fit

$$\chi^2 \equiv \sum_i \frac{(x^i_{pred} - x^i_{meas})^2}{\sigma_i^2}$$

Suppose our prediction is a relationship:

$$y_i = a x_i + b \qquad \text{(Measure "y")}$$

$$\chi^2 = \sum_i \frac{(a x_i + b - y_i)^2}{\sigma_i^2}$$

$$\frac{\partial \chi^2}{\partial a} = \sum_i \frac{2(a x_i + b - y_i) \cdot x_i}{\sigma_i^2} = 0$$

$$\frac{\partial \chi^2}{\partial b} = \sum_i \frac{2(a x_i + b - y_i)}{\sigma_i^2} = 0$$

$$\langle ? \rangle \equiv \sum_i \frac{?}{\sigma_i^2} \Big/ \sum_i \frac{1}{\sigma_i^2}$$

$$a \langle x^2 \rangle + b \langle x \rangle - \langle xy \rangle = 0$$

$$a \langle x \rangle + bN - \langle y \rangle = 0$$

$$Na\langle x^2\rangle + b\langle x\rangle N - \langle xy\rangle N = 0$$

$$a\langle x\rangle^2 + b\langle x\rangle N - \langle x\rangle\langle y\rangle = 0$$

$$a\left(\langle x^2\rangle - \langle x\rangle^2\right) = \langle xy\rangle - \langle x\rangle\langle y\rangle$$

$$\boxed{a = \frac{N\langle xy\rangle - \langle x\rangle\langle y\rangle}{N\langle x^2\rangle - \langle x\rangle^2}}$$

$$a\langle x^2\rangle\langle x\rangle + b\langle x\rangle^2 - \langle xy\rangle\langle x\rangle = 0$$

$$a\langle x\rangle\langle x^2\rangle + bN\langle x^2\rangle - \langle y\rangle\langle x^2\rangle = 0$$

$$b\left(N\langle x^2\rangle - \langle x\rangle^2\right) = \langle x^2\rangle\langle x\rangle - \langle xy\rangle\langle x\rangle$$

$$\boxed{b = \frac{\langle y\rangle\langle x^2\rangle - \langle xy\rangle\langle x\rangle}{N\langle x^2\rangle - \langle x^2\rangle}}$$

Central Limit Theorem:

1) We derived an example:
   Poisson Distribution ⟶ Gaussian Distribution
   for large N.

2) In scipy, we saw that

   many random numbers in $[0, 1]$ lead
   to Gaussian distribution

Generally CLT says arithmetic mean of
independent random variables with finite
variance converges to a Gaussian distribution
in limit of large N.

In case of Poisson we saw this convergence
can be *very fast*. $(N \sim 10)$.

In practice: this means the most
likely distribution we will encounter
is ... the Gaussian distribution.

Therefore; when we know no-better
we assume our PDF is Gaussian...

χ^2 Intro Lecture

χ^2 Tests

Want to understand how likely a set of data:

$$\{x_1 \pm \sigma_1, x_2 \pm \sigma_2, x_3 \pm \sigma_3, \dots\} \equiv x_i \pm \sigma_i$$

is the result of a corresponding theoretical prediction:

$$\{y_1, y_2, \dots, y_n\} \equiv y_i$$

Assume Gaussian uncertainties on x, then probability of one measurement is just a Gaussian PDF

$$P_i = \frac{1}{\sqrt{2\pi}\,\sigma_i} \exp\left(-\frac{1}{2}\frac{(y_i - x_i)^2}{\sigma_i^2}\right)$$

The probability for the complete set of n measurements is called the Likelihood and is just the probability

$$\mathcal{L} = \prod_i P_i = \prod_i \frac{1}{\sqrt{2\pi}\sigma_i} \exp\left(-\frac{1}{2}\frac{(y_i - x_i)^2}{\sigma_i^2}\right)$$

(In fact likelihood (and what follows) is quite a bit more general, we could have another PDF in some circumstances, and a different Likelihood function)

Best Estimates for X:

Suppose we make n measurements "drawn from a Gaussian distribution"

$$x_1, x_2, x_3, \ldots, x_n$$

What is our best estimate for the "true value" of x?

For a "true value" of X and an uncertainty of each measurement σ, the probability of measuring within Δx of x_i is

$$P(x_1) = \frac{1}{\sqrt{2\pi}\,\sigma} \exp\left(-\frac{(x_1 - X)^2}{2\sigma^2}\right) \Delta x$$

The combined probability of our whole series is

$$P = P(x_1) \cdot P(x_2) \cdot P(x_3) \cdots P(x_n)$$

$$= \left(\frac{\Delta x}{\sqrt{2\pi}\,\sigma}\right)^N \exp\left(-\frac{(x_1 - X)^2}{2\sigma^2}\right) \cdot \exp\left(-\frac{(x_2 - X)^2}{2\sigma^2}\right) \cdots$$

$$= \left(\frac{\Delta x}{\sqrt{2\pi}\,\sigma}\right)^N \exp\left(-\frac{\sum_i (x_i - X)^2}{2\sigma^2}\right)$$

- Since sums are easy than products, and log is monotonically increasing

$\log \mathcal{L}$

- Since we physicists think about minimization (as opposed to maximization)

$-\log \mathcal{L}$

- Since there is an annoying factor of $\frac{1}{2}$ in exponent, scale

$-2\log \mathcal{L}$

Let's calculate it:

$$-2\log \mathcal{L} = -2\sum_i \log\left(\frac{1}{2\pi\sigma_i}\right) + \sum_i \frac{(y_i - x_i)^2}{\sigma_i^2}$$

Does not depend on x_i or y_i, only a constant depending on precision of experiment

$\equiv \chi^2$

$$\chi^2 = \sum_i \frac{(y_i - x_i)^2}{\sigma_i^2} \quad \left(= -2\log\mathcal{L} + \text{const}\right)$$

SMALL $\chi^2 \Rightarrow$ Large Probability data and prediction agree

(*) For counting exp $x_i \to N_i$ $\sigma_i^2 \to N_i$

$$\chi^2 = \frac{(m - x_1)^2}{\sigma_1^2} + \frac{(m - x_2)^2}{\sigma_2}$$

$$\frac{d\chi^2}{dm} = 2\,\frac{(m - x_1)}{\sigma_1^2}$$

$$\chi^2 = \sum \frac{(m - x_i)^2}{\sigma_i'^2}$$

$$\frac{d\chi^2}{dm} = \sum_i \frac{2(m - x_i)}{\sigma_i'^2}$$

$$\frac{d^2\chi^2}{dm^2} = 2 \sum_i \frac{1}{\sigma_i'^2}$$

$$\frac{1}{2}\,\frac{\partial^2\chi^2}{dm^2} = \sum_i \frac{1}{\sigma_i'^2} = \frac{1}{\sigma_m^2}$$

$$= \frac{N}{\sigma^2} = \frac{1}{\sigma_m}$$

$$\sigma_m^2 = \frac{\sigma^2}{N}$$

Chapters 1 – 3

Uncertainties

→ Scientist believe (or imagine) they are perfect!

Error ≠ Mistake

Error = Uncertainty

→ Why are uncertainties useful?

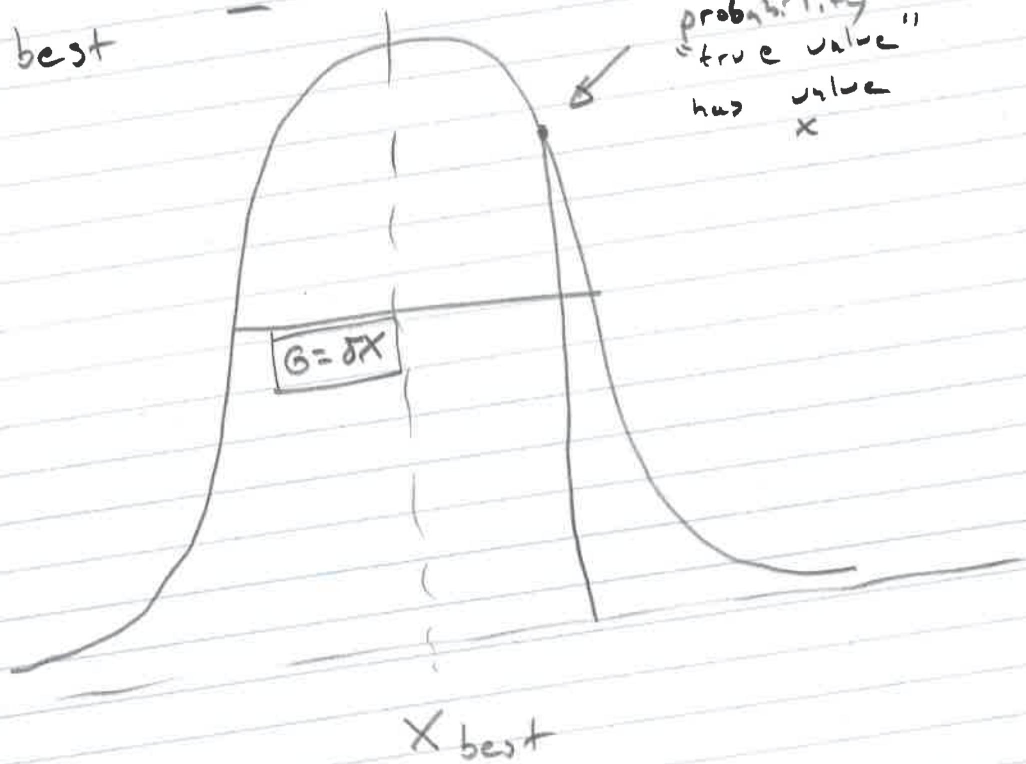
→ Report the precision of the measurement.

→ Measurement can be compared to other measurements

→ Measurement can be compared to theory predictions.

Gaussian Uncertainties:

$X_{best} \pm \delta X$



Conventions

$X_{best} \pm \delta X$ means PDF is Gaussian with

Why?

1) CLT.
2) Simple: 2 parameters: mean and sigma.
3) Math is easy... as we will see, extracting best values of x and σ is just

— mean
— rms

Error Function

Suppose we measure the speed of light to be:

$$3.3 \pm 0.2 \times 10^8 \ m/s$$

How consistent is this with the generally accepted value

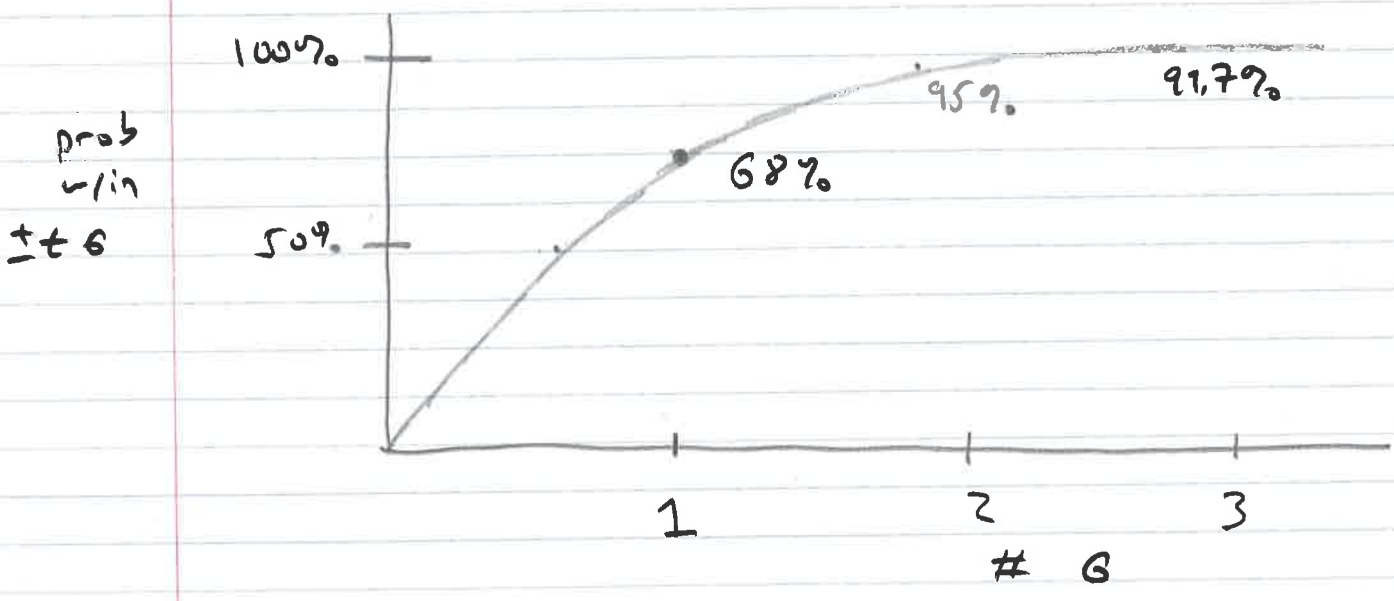
$$3.0 \times 10^8 \ m/s$$

First answer: $\frac{3.3 - 3.0}{0.2} = \boxed{1.5 \text{ sigma}}$

But what if we want to know probability enclosed within $\pm t \cdot \sigma$.

$$\boxed{\mathrm{erf}(t) = \frac{1}{\sqrt{2\pi}} \int_{-t}^{t} \exp\left(-\frac{x^2}{2}\right)}$$

No analytic solution, so we just define this function and tabulate it...



HEP uses 5 σ as discovery threshold

$\boxed{99,99994\ \%}$

Our choice to report errors as 1 σ for Gaussian means
"true value within uncertainty 68% of time"