Lecture 18: Maximum Likelihood Fitting

\[ L(\lambda | t_1, t_2, \ldots, t_n) = \prod_{i=1}^{n} f(t_i) \]

\[ = \prod_{i=1}^{n} \lambda e^{-\lambda t_i} \]

\[ = \lambda^n \cdot e^{-\lambda \sum_{i=1}^{N} t_i} \]
Before Spring Break...

• Introduction to “fitting”
• Compare data ↔ physical model
• $\chi^2$ minimization
  – Gaussian-distributed uncertainties
  \[
  \chi^2 = \sum_{i=1}^{n} \frac{(f(x_i) - y_i)^2}{\sigma^2_{y_i}}
  \]
  – Best fit parameters are those that minimize this function
  – Many ways to get there!
A07: Linearization

- We know how to do linear least-squares fitting (fit.cpp)
- But our function isn't a straight line:

\[ F(v) = C v^\beta \]
A07: Linearization

- Its easy to linearize the model:

\[ F(v) = C v^\beta \]

\[ \ln F = \ln C v^\beta \]

\[ = \ln C + \beta \ln v \]
A07: Linearization

- What about error bars?
- Use standard error propagation equation:

\[ \xi = \ln F \]

\[ \sigma_\xi = \left( \frac{d\xi}{dF} \right) \sigma_F = \frac{\sigma_F}{F} \]
Comments on Fit Output

Ignoring standard deviations

\[ a = 114.281062 \quad \text{uncertainty: 10.787334} \]
\[ b = 31.476746 \quad \text{uncertainty: 1.000535} \]
\[ \text{chi-squared: } 4052.544699 \]
\[ \text{goodness-of-fit: } 1.000000 \]

Including standard deviations

\[ a = 119.496706 \quad \text{uncertainty: 7.567595} \]
\[ b = 30.697872 \quad \text{uncertainty: 1.034080} \]
\[ \text{chi-squared: } 10.946476 \]
\[ \text{goodness-of-fit: } 0.204750 \]
Maximum Likelihood Method
\( \chi^2 \) method requires Gaussian Uncertainty

- Simplest case: straight line fit to data with *normally distributed* (Gaussian) uncertainties.

- What does this mean?
  - Measure quantity \( x \), \( N \) times.
  - As \( N \to \text{large} \), parent probability distribution \( \to \) Gaussian

\[
f(x) = \frac{1}{\sqrt{2\pi\sigma}} e^{-\frac{(x-x_0)^2}{2\sigma^2}}
\]
What if uncertainties Non-Gaussian?

- “Maximum Likelihood Method”
  - Unbinned fitting technique
  - $\chi^2$ method is actually special case of MLM

- Define the “Likelihood Function”

\[
\mathcal{L}(a_1, a_2, \ldots, a_m) = \prod_{i=1}^{N} P_i(x_i; a_1, a_2, \ldots, a_m)
\]
What if uncertainties Non-Gaussian?

- “Maximum Likelihood Method”
  - Unbinned fitting technique
  - $\chi^2$ method is actually special case of MLM

- Define the “Likelihood Function”

$$
\mathcal{L}(a_1, a_2, \ldots, a_m) = \prod_{i=1}^{N} P_i(x_i; a_1, a_2, \ldots, a_m)
$$

- Number of measurements
- Quantity being measured
- Product
- Probability density function evaluated for $i^{th}$ data point
- Fit parameters
Example: Muon Decay

- Elementary particles exhibit exponential decay.
- $\tau$ is “mean lifetime” of the particle.
- How do we extract $\tau$?
\( \chi^2 \) method:

- Sort lifetimes of individual particles into bins of width \( \Delta t \).
- Assign \( \sqrt{n} \) uncertainty to the contents of each bin.
\( \chi^2 \) method:

- Sort lifetimes of individual particles into bins of width \( \Delta t \).
- Assign \( \sqrt{n} \) uncertainty to the contents of each bin.
- Linearize (take log)
- Find \( \tau \) which minimizes \( \chi^2 \).
\( \chi^2 \) method:

- Sort lifetimes of individual particles into bins of width \( \Delta t \).
- Assign \( \sqrt{n} \) uncertainty to the contents of each bin.
- Linearize (take log)
- Find \( \tau \) which minimizes \( \chi^2 \).
- Any potential problems?
Muon Decay: Likelihood method

For muon decay...

\[ P_i = A e^{-t_i/\tau} \]

where \( A \) is a normalization constant subject to

\[
\int_{t_1}^{t_2} P_i dt_i = A \int_{t_1}^{t_2} e^{-t_i/\tau} dt_i = 1
\]

If \( t_1 = 0, \ t_2 = \infty, \)

\[ A = \frac{1}{\tau} \]

otherwise

\[ A = \frac{1}{\tau \left[ e^{-t_1/\tau} - e^{-t_2/\tau} \right]} \]
Muon Decay: Likelihood method

- Practically speaking, since the $P_i < 1$, $L$ can be very small.

- More commonly what is maximized is the log of the likelihood function...

$$\mathcal{M} = \ln \mathcal{L}$$

$$= \ln \left[ \prod_{i=1}^{N} (A e^{-t_i/\tau}) \right]$$

$$= \sum_{i=1}^{N} \left[ \ln A - \frac{t_i}{\tau} \right]$$
Wrap-up

• Today's lab: Apply grid-search method to find value of parameter $\tau$ which maximizes log-likelihood.

• Uncertainties: What change in $\tau$ changes $M$ by 1/2?

• Goodness of fit (disadvantage of direct ML method): Requires other methods, e.g. large simulated (Monte Carlo) data sets to derive!