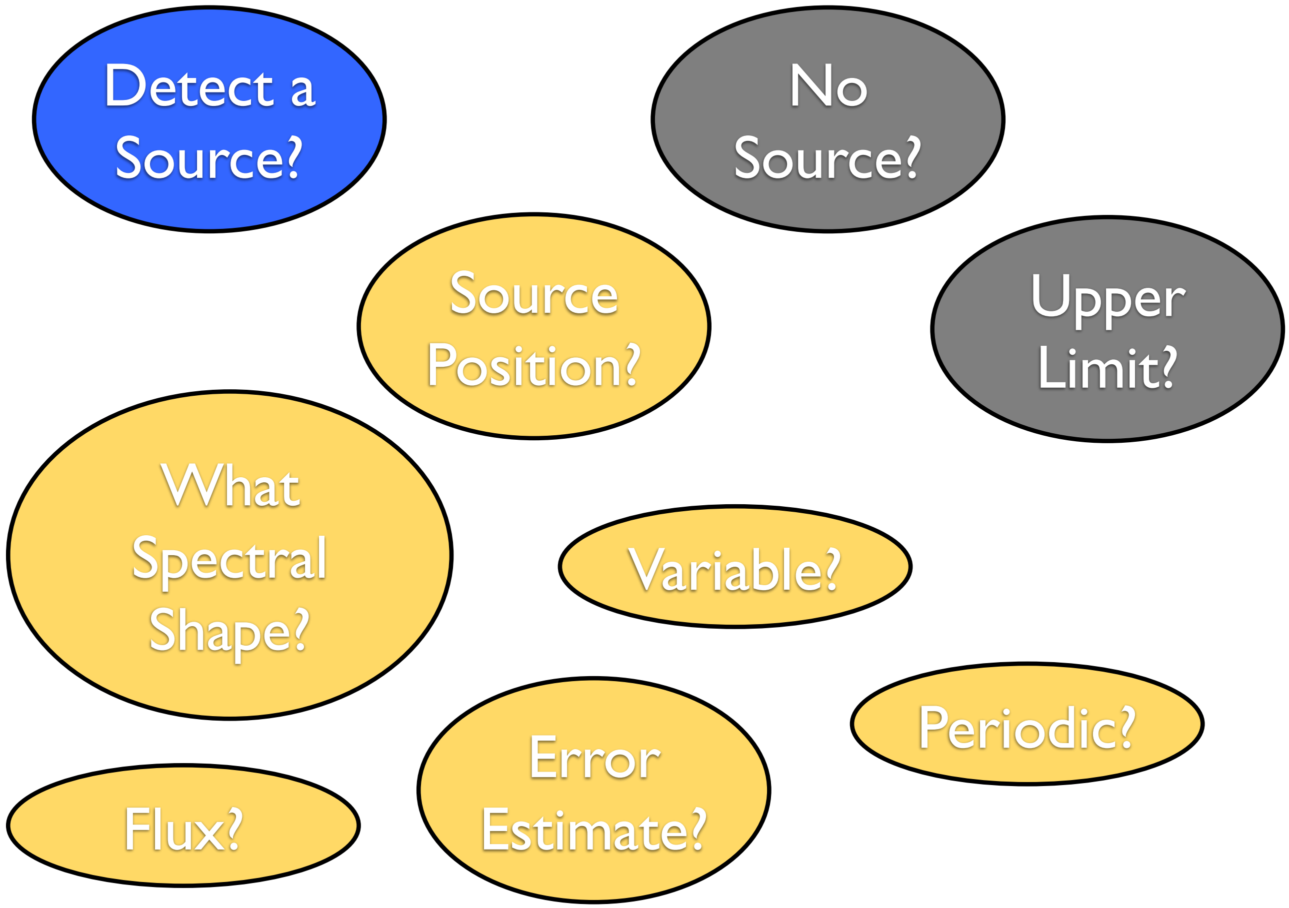
Notes on Maximum Likelihood in LAT

What can I infer from my observation?



Measurements in γ-ray astronomy

- **Hypothesis testing** — Is a source significantly detected?
  - **Parameter estimation** — If so, what is its flux?
  - **Hypothesis testing / Parameter estimation** — If not, what upper limit on the flux?
- **Hypothesis testing** — What kind of spectrum does it have?
  - **Parameter estimation** — What is its spectral index?
- **Parameter estimation** — What is its location in the sky?
- **Hypothesis testing / Parameter estimation** — What are the errors on these values?
- **Hypothesis testing** — Is the source variable?

~~Good things~~ about maximum likelihood

Cautions

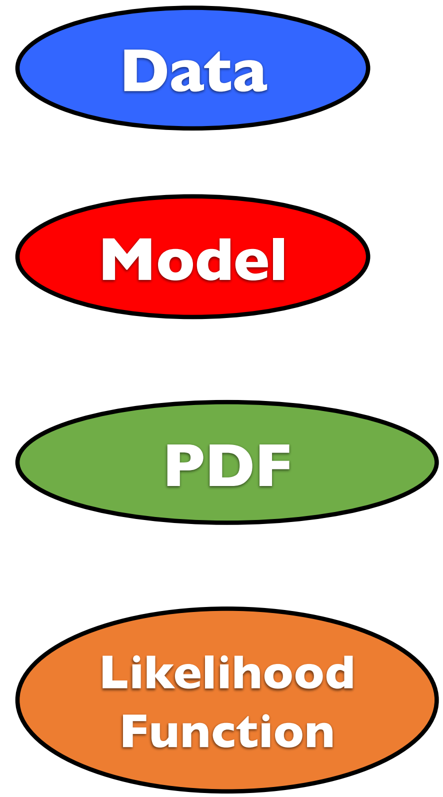
- General framework for statistical questions.
- Unbiased, minimum variance estimate as sample size increases.
- Asymptotically Gaussian: allows evaluation of confidence bounds & hypothesis testing.
- Well studied in the literature.
- Starting point for Bayesian analysis.

- Only answers the question asked.
- Be aware of small number regimes and departure from Gaussian assumption for confidence bounds and hypothesis testing
- Starting point for Bayesian analysis.

Maximum likelihood technique

Given a set of observed data

- produce a model that *accurately* describes the data, including parameters that we wish to estimate,
- derive the probability (density) for the data given the model (probability density function, PDF),
- treat this as a function of the model parameters (likelihood function), and
- maximize the likelihood with respect to the parameters - ML estimation.



ML estimation (MLE)

- Parameters can be estimated by maximizing likelihood. Easier to work with log-likelihood:

$$\ln \mathcal{L}(\Theta) = \ln \mathcal{L}(\Theta|X) = \sum_i \ln P(x_i|\Theta)$$

- Estimates of parameters $\{\hat{\theta}_k\}$ from solving simultaneous equations: $\left.\frac{\partial \ln \mathcal{L}}{\partial \theta_j}\right|_{\{\hat{\theta}_k\}} = 0$

- For one parameter, if we have: $\mathcal{L}(\theta) \sim e^{-\frac{(\theta-\hat{\theta})^2}{2\sigma_\theta^2}}$

  then: $\left.\frac{\partial^2 \ln \mathcal{L}}{\partial \theta^2}\right|_{\hat{\theta}} = -\frac{1}{\sigma_\theta^2}$

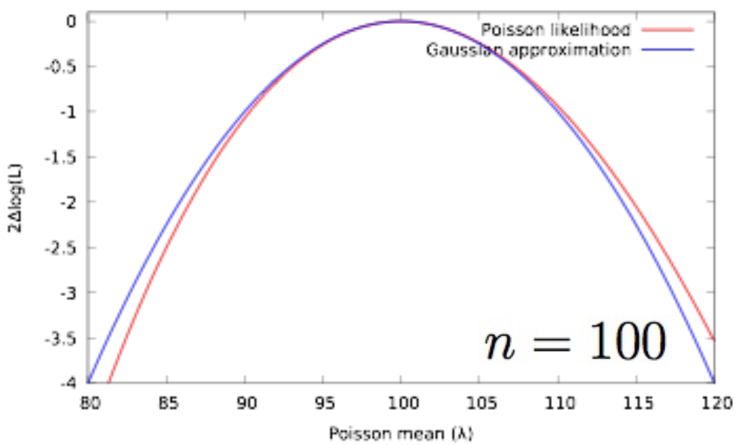
  Gaussian approximation

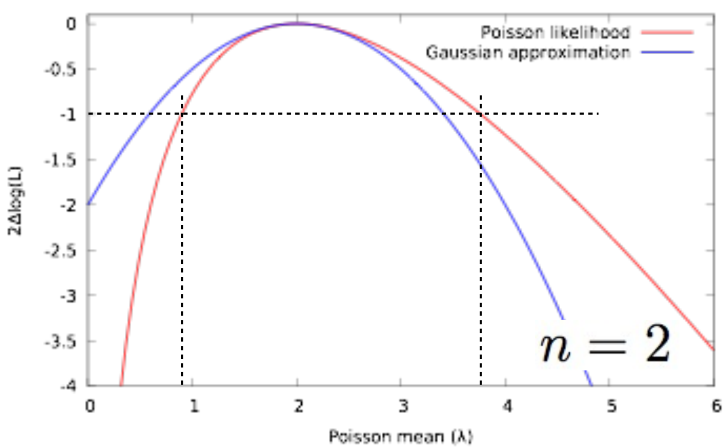
  2nd derivative is related to "errors"

Binned Likelihood

- Ideally, we would look up the exact response for each individual photon and add up each individual log likelihood
- That can add up to a lot of calculation for longer intervals. Instead group photons in similar bins in inclination angle and energy
  - The spatial and energy bins have to be small enough that the response is not very different for photons in the bin
- Numbers of total photons may be large, but numbers in any bin may be small

Log-likelihood profile and errors





Large number of events – Gaussian approximation reasonably accurate

$$\sigma_\lambda^2 = n$$

Log-likelihood profile provides a more accurate estimate for small number of events

$$2 \ln \mathcal{L}(\lambda) = 2 \ln \mathcal{L}(\hat{\lambda}) - 1$$

n = 100; $\hat{\lambda} = 100.0^{+10.33}_{-9.67}$

Log-likelihood profile provides a better error estimate

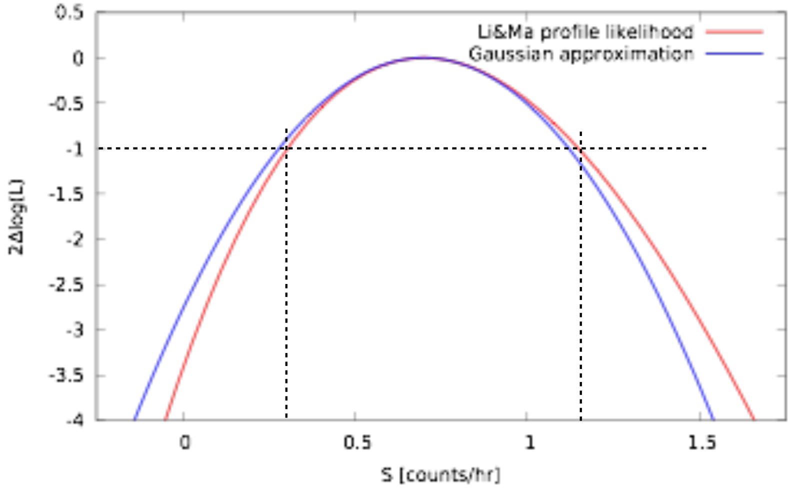
n = 2; $\hat{\lambda} = 2.0^{+1.77}_{-1.10}$

Profile likelihood

Confidence regions with nuisance parameters
Rolke, et al., NIM A, 551, 493 (2005)

- Often we are either concerned only with the one parameter, or wish to treat the multiple parameters separately (ignore covariance).
- Produce "profile log-likelihood" curve, a function of only one parameter (at a time), maximized over all others.
- LRT says this should behave as $\chi^2(1)$.
- Define confidence region using this function exactly as before.

Example of profile likelihood



Rate $\hat{S} = 0.7^{+0.45}_{-0.39}\,\mathrm{hr}^{-1}$

This is not a significant result, so we would usually not claim a detection. Provide an upper limit instead.

- Use simple On/Off counting example

$$n_{off} = 24$$
$$n_{on} = 15$$
$$\alpha = 1/3$$
$$T = 10.0\,\mathrm{hr}$$

- Giving:

$$\hat{S} = 0.7\,\mathrm{hr}^{-1}$$
$$\sigma_S = 0.42\,\mathrm{hr}^{-1}$$
$$TS = 3.43$$
$$\sigma = 1.85$$

For next time

- Homework
  - Pick a source or topic you want to explore this summer. Google form here.
  - Try making a profile log likelihood plot for a parameter in the 3C 279 example analysis.
  - Bonus example on source finding will be posted later today. We will come to this next week and discuss hypothesis testing.

- Next time
  - Special case: Light curves , or running a series of likelihood analyses