Statistics and parameter estimation in cosmology

Observational Cosmology

NASSP course

Bruce Bassett

Resources

- Wikipeadia.com + lots on the web

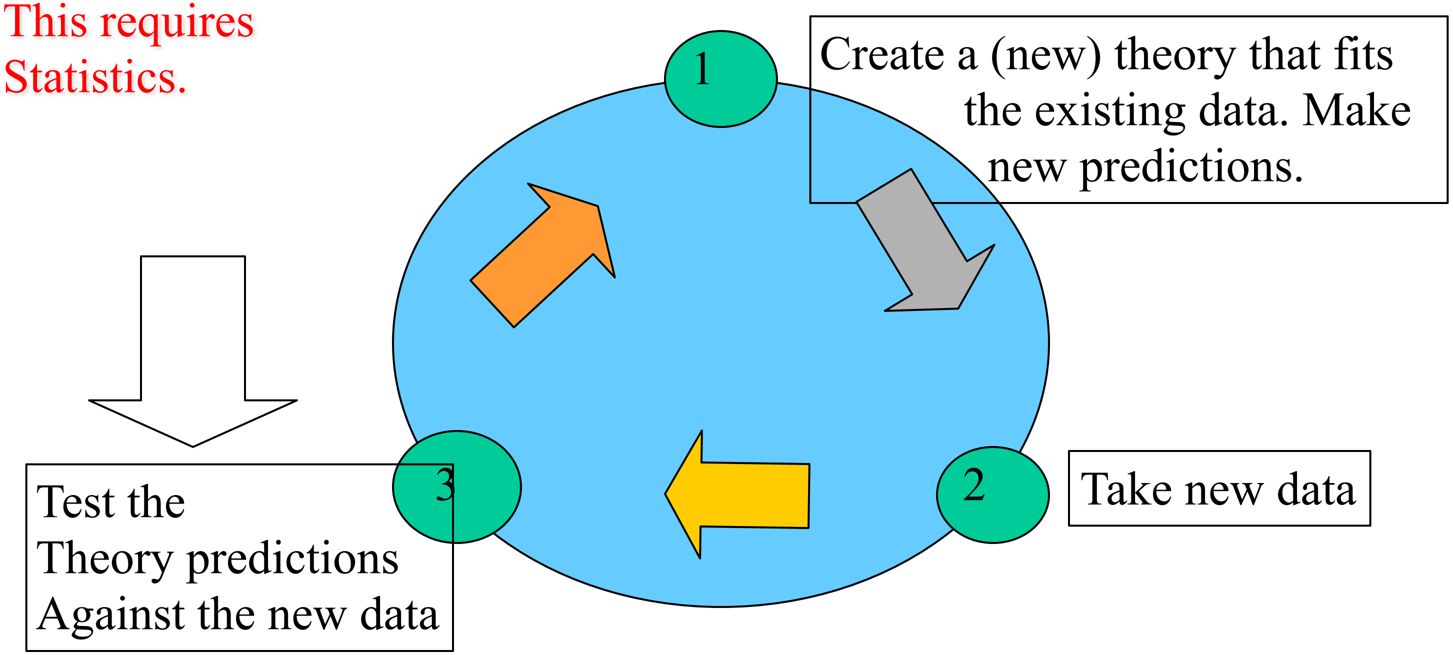
- D. J. C. MacKay, *Information theory, inference, and learning algorithms* Cambridge (2003). A good how-to book for MCMC, neural networks, data compression.

- Available on the web at: http://www.inference.phy.cam.ac.uk/mackay/itila/book.html

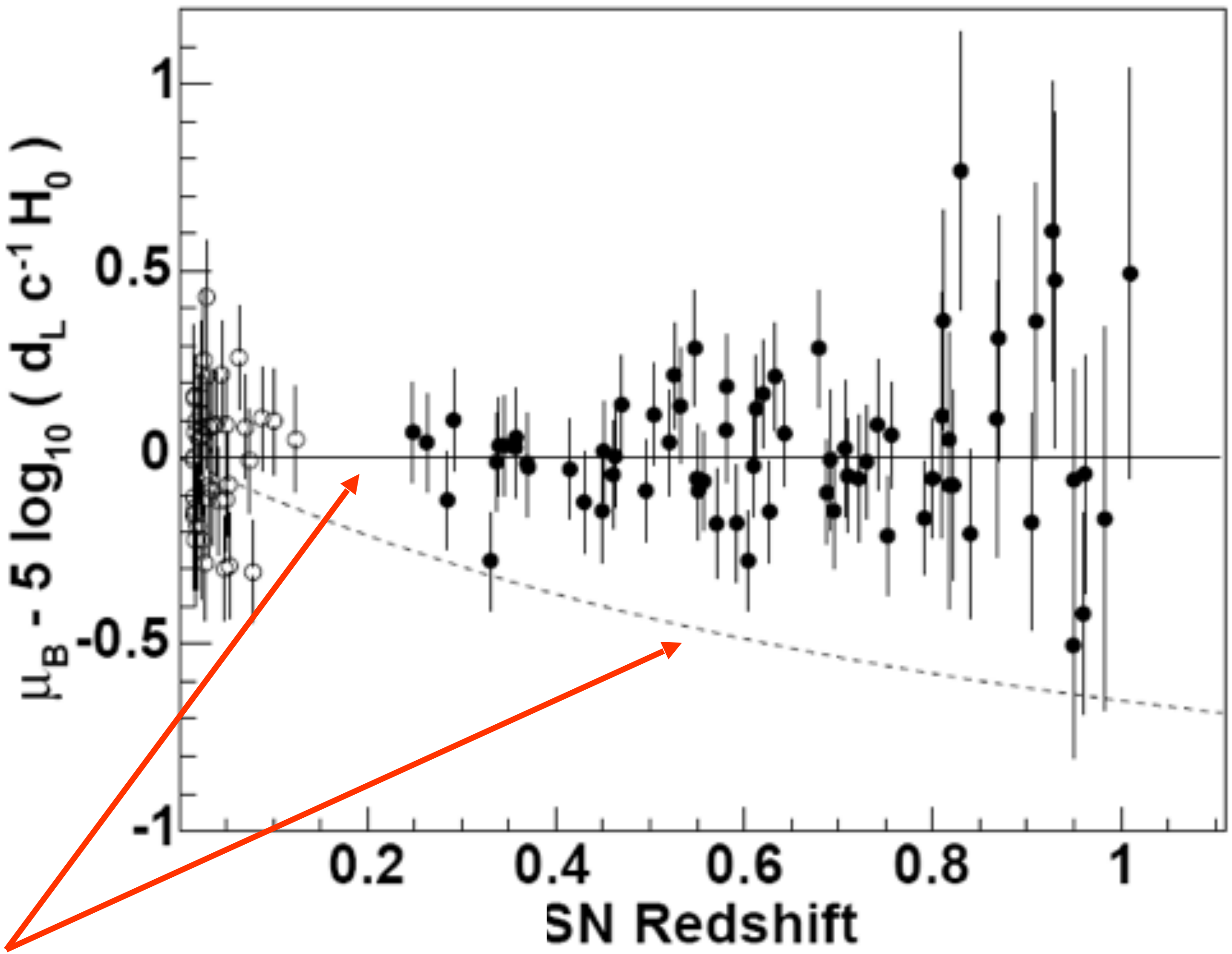
What's the point?

- A crude characterisation of science is that it is the full circle:

This requires Statistics.

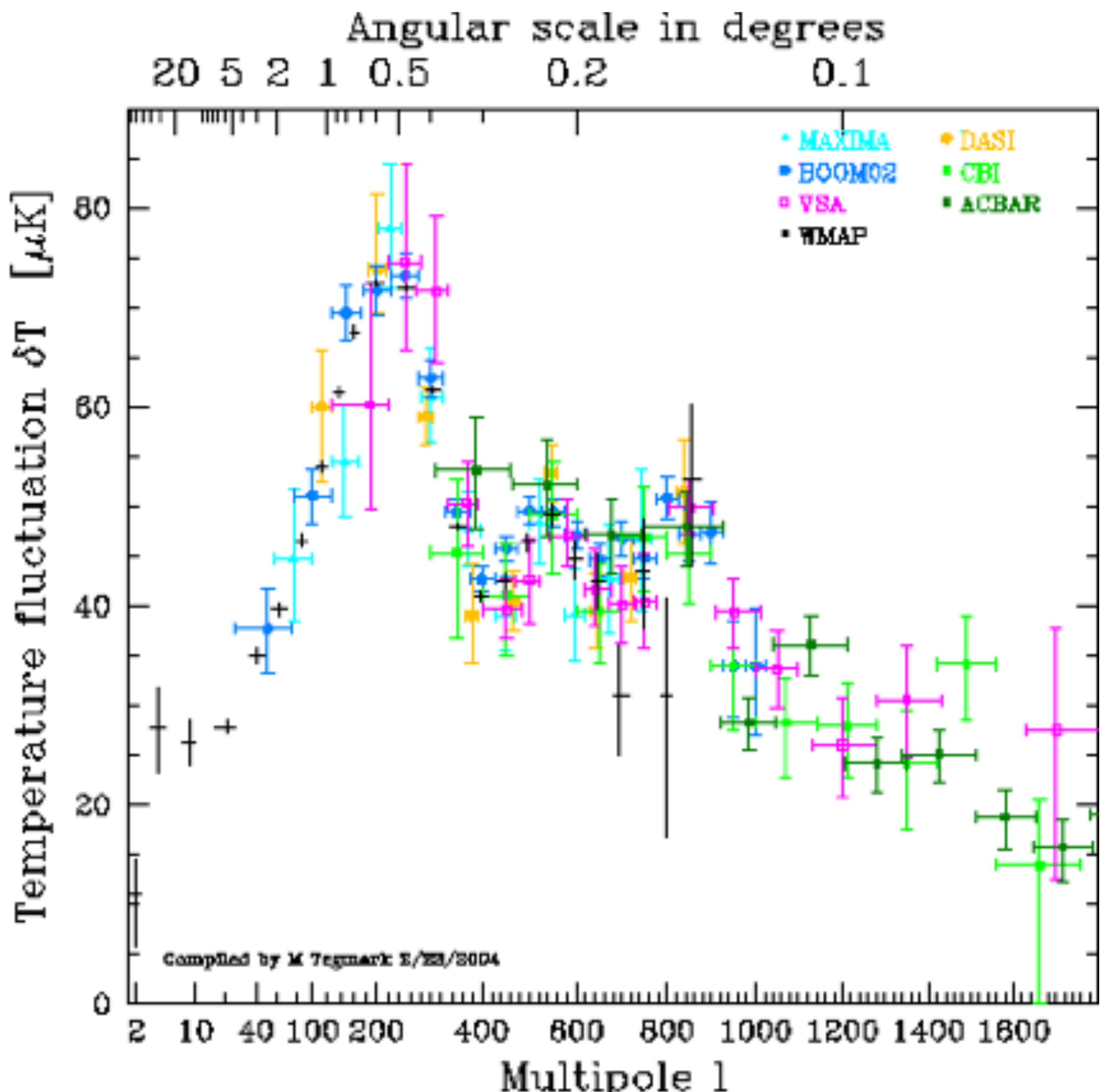


The problem…

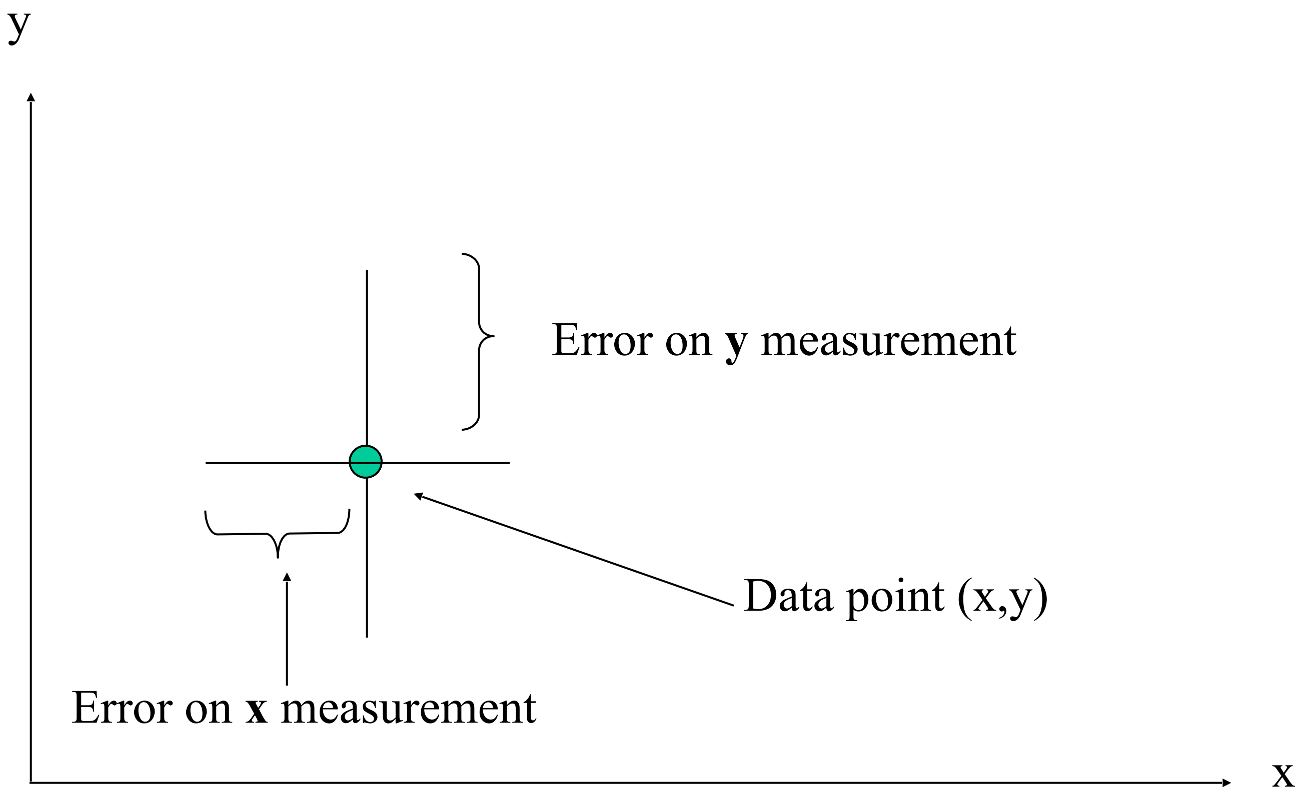


The solid line
Is clearly a better fit
To the data…but how do we quantify this belief?

…and how do we combine multiple datasets?



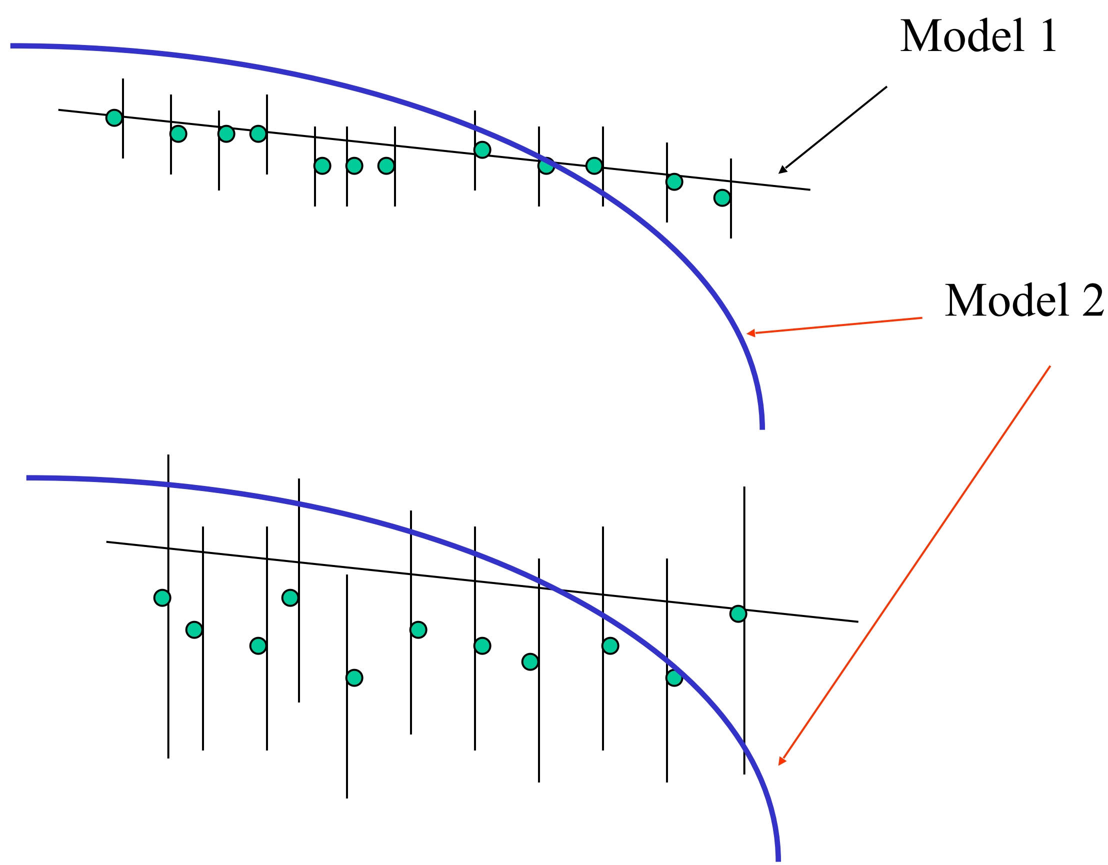
What are error bars?



error bar sizes

Here model 2 is ruled out by the data at high confidence.

Here model 2 is not ruled out by the data.



The χ^2 statistic

- A useful measure of how well data is fit by a theoretical curve (which may depend on some parameters, θ_ι) is the quantity:

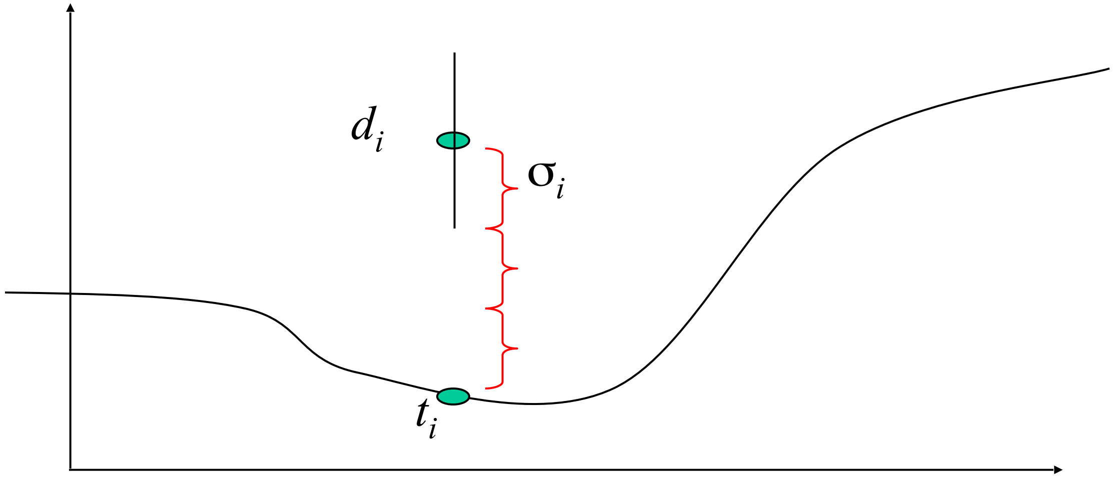
$$\chi^2 = \sum_{i=1} \left(\frac{d_i(x_i) - t(x_i, \theta)}{\sigma_i} \right)^2$$

d_i is the data
t_i is the theoretical prediction at that point.
The sum is over data points.

- This is the sum of the number of standard deviations (σ_i) the data is away from the theoretical curve.

What is the χ^2?

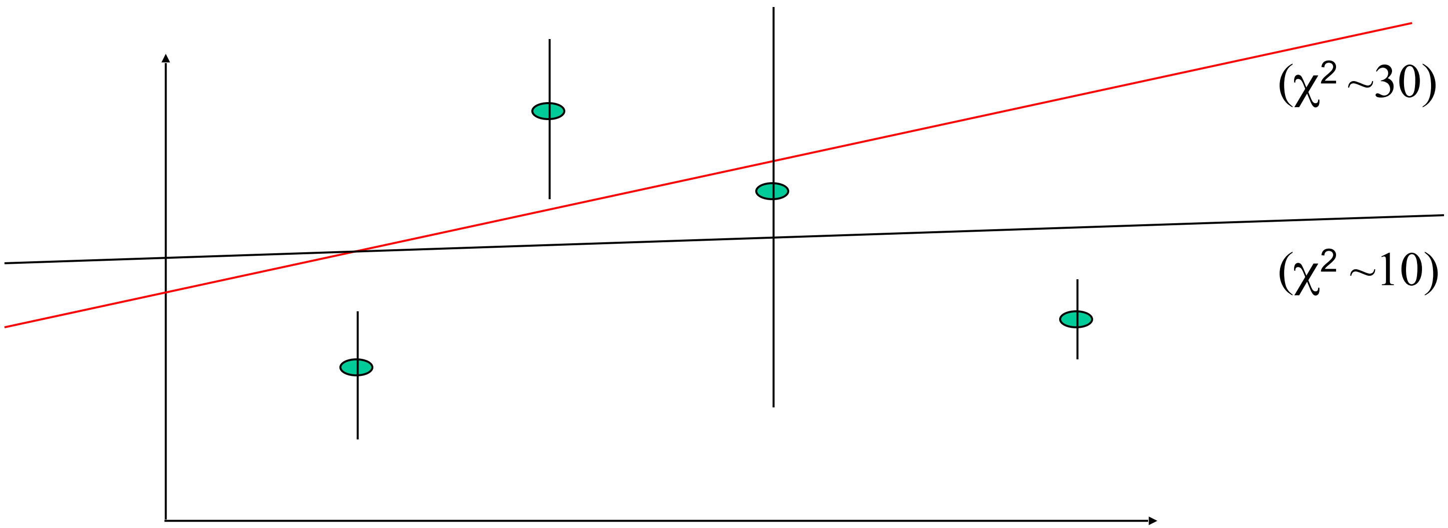
$$\chi^2(\theta_\alpha) \equiv \sum_i \frac{(d_i - t_i(\theta_\alpha))^2}{\sigma_i^2}$$



- This point has $\chi^2 \sim 9$

$$\chi^2(\theta_\alpha) \equiv \sum_i \frac{(d_i - t_i(\theta_\alpha))^2}{\sigma_i^2}$$

- χ^2 is just the sum over the data points, of the number of standard deviations the data is from the model:



χ^2 continued...

- **Exercise 1:** assuming a linear model:

$$t(x,\theta_i) = \theta_1 x + \theta_2$$

Analytically compute the best-fit values of the parameters given a set of *N* data-points with given errors, σ_i, using the χ^2 statistic.

Parameter estimation - I

- A very common problem in science is that we want to find the parameter values that best fit given observational data.
- For example: what is the period of a binary star, the mass of the electron, the value of the Hubble constant or the size of the cosmological constant?

Parameter estimation - II

- How can we estimate the best-fitting parameters?
- Since any measurement has finite accuracy, this is always a statistical problem. We want the most-likely values of the parameters.
- We can estimate this by minimising χ^2 while we vary the parameters of the theory. We already did it in the case of a linear model.

To estimate parameters…

$$\frac{\partial \chi^2}{\partial \theta_\alpha} = 0 \ldots. \alpha = 1 \ldots n$$

Parameter estimation - III

- If the model depends nonlinearly on the parameters we usually cannot solve it analytically and must use numerical methods.
- This is a well-studied branch of computational mathematics.
- We want to find the global minimum of χ^2 (there are often lots of local minima which we are not interested in)

Parameter estimation - IV

- One of the easiest ways to minimise χ^2 when there are few parameters (less than 4 or 5) is the *grid method*.
- Consider the case of a single parameter, θ, for simplicity. We split up the allowed range of θ into n equal regions, with grid points labeled θ_i.
- We then compute χ^2 at each point in the region and then choose the smallest χ^2_i . The corresponding θ_i is approximately the best-fit parameter value. Obviously the method can be refined by increasing n.
- Grid method fails for large number of parameters (why?). Stochastic/Monte Carlo methods are much better in this case (e.g. MCMC, simulated annealing).