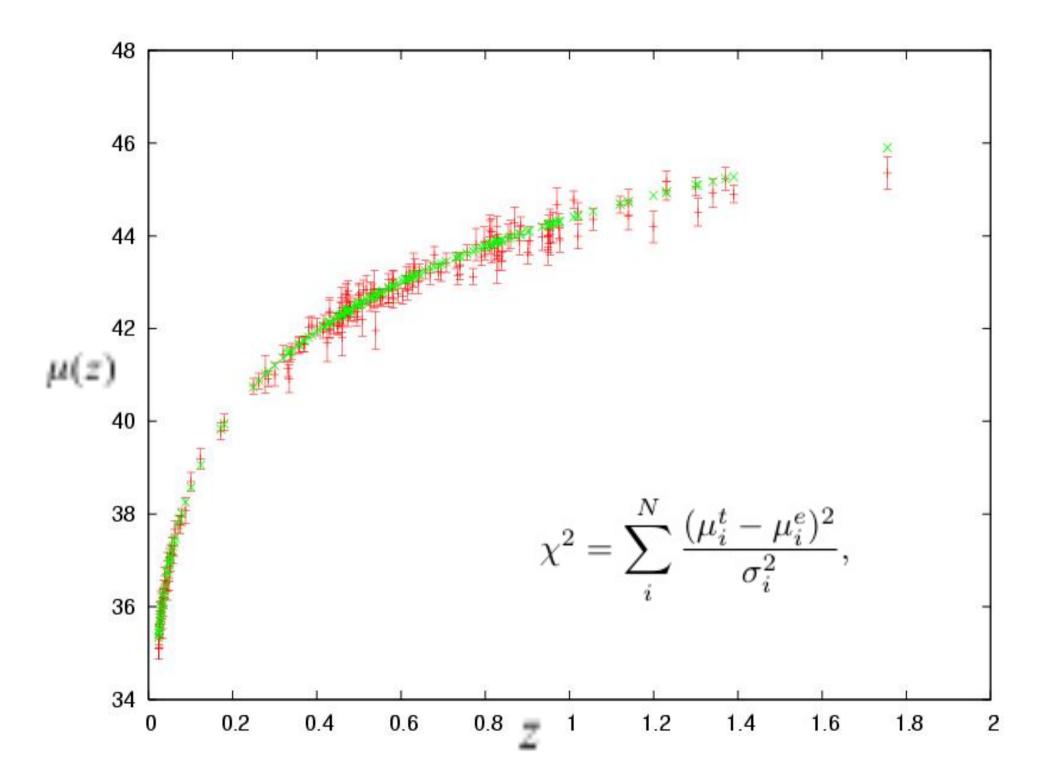
Crossing Statistic

Arman Shafieloo Institute for the Early Universe, Ewha University

4th KIAS Workshop on Cosmology and Structure Formation November 2010, Seoul-Korea

Minimization of reduced Chi square or effective Chi square is the most common approach in cosmology (and many other fields of science) to do parameter estimation and also being used in model selection.

$$\chi^2 = \sum_i^N \frac{(\mu_i^t - \mu_i^e)^2}{\sigma_i^2},$$

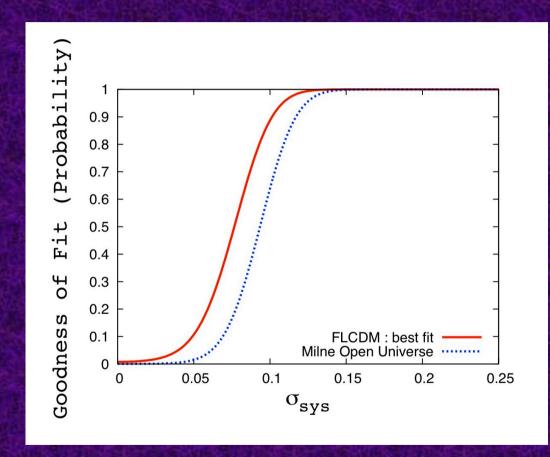


Consistency of a model and the data:

- Frequentist Approach:
- Assuming a proposed model, the probability of the observed data must not be insignificant.
- Bayesian Approach:
- Priors and simplicity of the proposed model also matters (in model comparison)

Chi square analysis plays a crucial role in calculation of the likelihood in both approaches

What if the actual size of the error bars are not known?



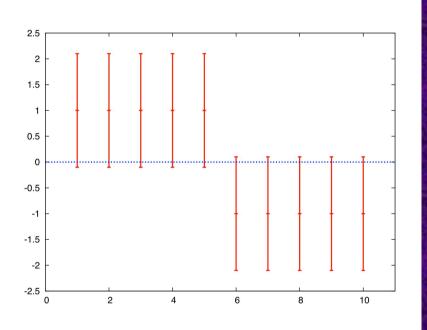
$$\chi^2 = \sum_{i}^{N} \frac{(\mu_i^t - \mu_i^e)^2}{\sigma_i^2},$$

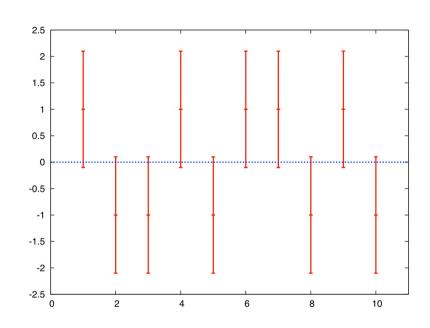
$$\sigma_i^2 = \sigma_{i(data)}^2 + \sigma_{(sys)}^2$$

$$P(\chi^2;N) = \frac{2^{-N/2}}{\Gamma(N/2)} \chi^{N-2} e^{-\chi^2/2}$$

$$Prob(\chi^2;N) = \int_{\chi^2}^{\infty} P(\chi^2;N) d\chi'^2.$$

Equal in being probable?!



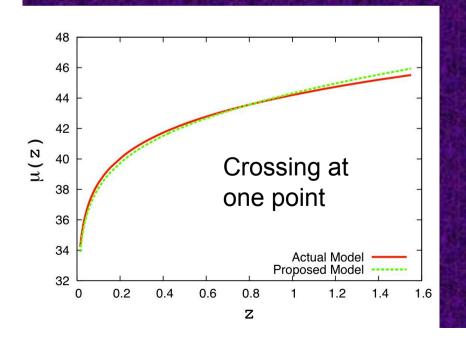


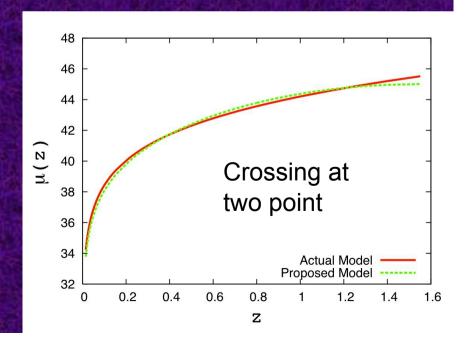
$$\chi^2 = \sum_{i}^{N} \frac{(\mu_i^t - \mu_i^e)^2}{\sigma_i^2},$$

Crossing Statistic

If a proposed model is different than the actual model, then they cross each other at one or two or three or ... N points.

A. Shafieloo, T. Clifton & P. Ferreira, 2010





One point Crossing: T1

- 1. Assume a model
- 2. Construct the normalized residuals
- 3. Finding the crossing point and calculating T1 by maximizing T(n1):
- 4. Comparing the results with Monte Carlo simulations.

$$q_i(z_i) = \frac{\mu_i(z_i) - \bar{\mu}(z_i)}{\sigma_i(z_i)}.$$

$$T(n_1) = Q_1(n_1)^2 + Q_2(n_1)^2,$$

$$Q_1(n_1) = \sum_{i=1}^{n_1} q_i(z_i)$$

$$Q_2(n_1) = \sum_{i=n_1+1}^{N} q_i(z_i),$$

Two points Crossing: T2

- 1-2.....
- 3. Finding the crossing points and calculating T2 by maximizing T(n1,n2):

$$q_i(z_i) = \frac{\mu_i(z_i) - \bar{\mu}(z_i)}{\sigma_i(z_i)}.$$

$$T(n_1, n_2) = Q_1(n_1, n_2)^2 + Q_2(n_1, n_2)^2 + Q_3(n_1, n_2)^2,$$

4. Comparing the results with Monte Carlo simulations.

And so on we can derive T3, T4,...

$$Q_1(n_1, n_2) = \sum_{i=1}^{n_1} q_i(z_i)$$

$$Q_2(n_1, n_2) = \sum_{i=n_1+1}^{n_2} q_i(z_i)$$

$$Q_3(n_1, n_2) = \sum_{i=n_1+1}^{N} q_i(z_i).$$

Important Features:

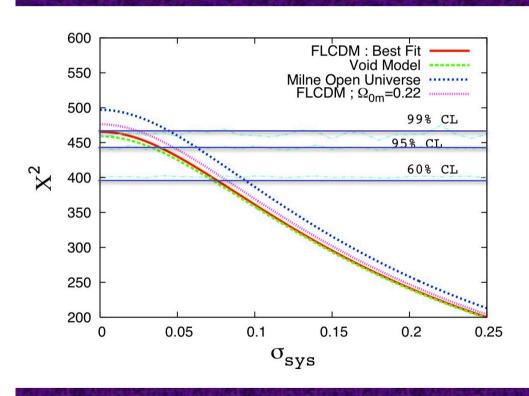
For N data points, the last mode of Crossing Statistic is T(N-1) which is identical to Chi Square Statistic

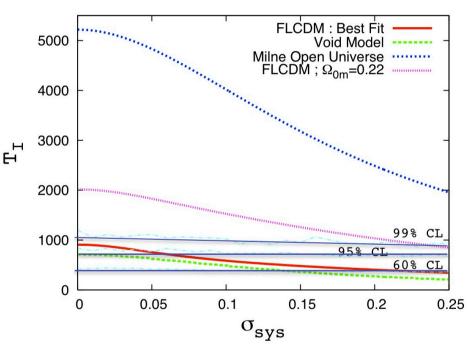
$$T_{N-1} = \sum_{i}^{N} (q_i)^2 = \chi^2$$

The zero mode of Crossing Statistic is similar to Median Statistic (Gott et al 2001)

$$T_0 = (\sum_i^N q_i)^2$$

Some Applications:

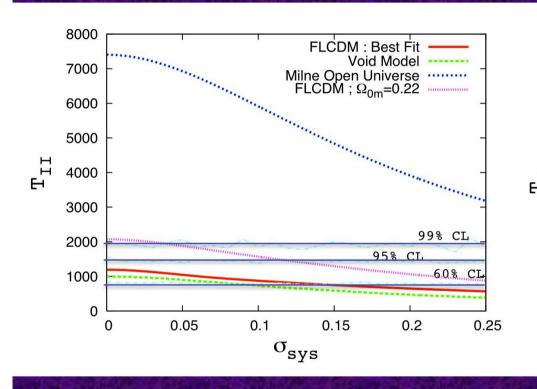


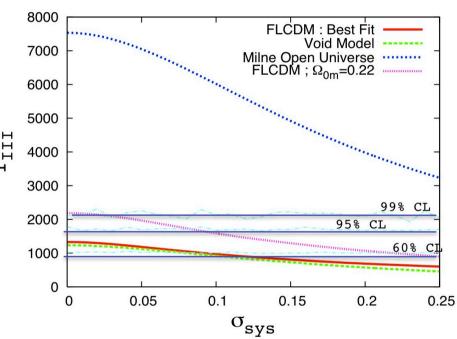


Constitution
Supernovae data

$$\sigma_i^2 = \sigma_{i(data)}^2 + \sigma_{(sys)}^2$$

Some Applications:





Constitution
Supernovae data

$$\sigma_i^2 = \sigma_{i(data)}^2 + \sigma_{(sys)}^2$$

Comparing Two Statistics

	T1	Chi Square
Ruling out by 99% CL	1% (Actual Model) 28.5% (Wrong Model)	1% (Actual Model) 1.9% (Wrong Model)
Ruling out by 99% CL Assuming extra (0.05) intrinsic dispersion	0.5% (Actual Model) 26.4% (Wrong Model)	0% (Actual Model) 0% (Wrong Model)

Actual Model: Flat LCDM with $\Omega_{0m}^{true} = 0.27$

 $\sigma_i^2 = \sigma_{i(data)}^2 + \sigma_{(sys)}^2 \qquad \text{Wrong Model: Flat LCDM with } \Omega_{0m}^{\textit{erroneous}} = 0.22$

Simulated data similar to Constitution compilation

Summary:

- Designed to be used in situations where the intrinsic dispersion of a data set is not well known (e.g supernovae data). Crossing statistic is in general less sensitive than x2 to the intrinsic dispersion of the data.
- Crossing Statistic can easily distinguish between different models in cases where the χ2 statistic fails.
- The last mode of Crossing Statistic is identical to χ2, so that one can consider it as a *generalization of χ2*. In fact we are extracting more information from the data by using Crossing Statistic.