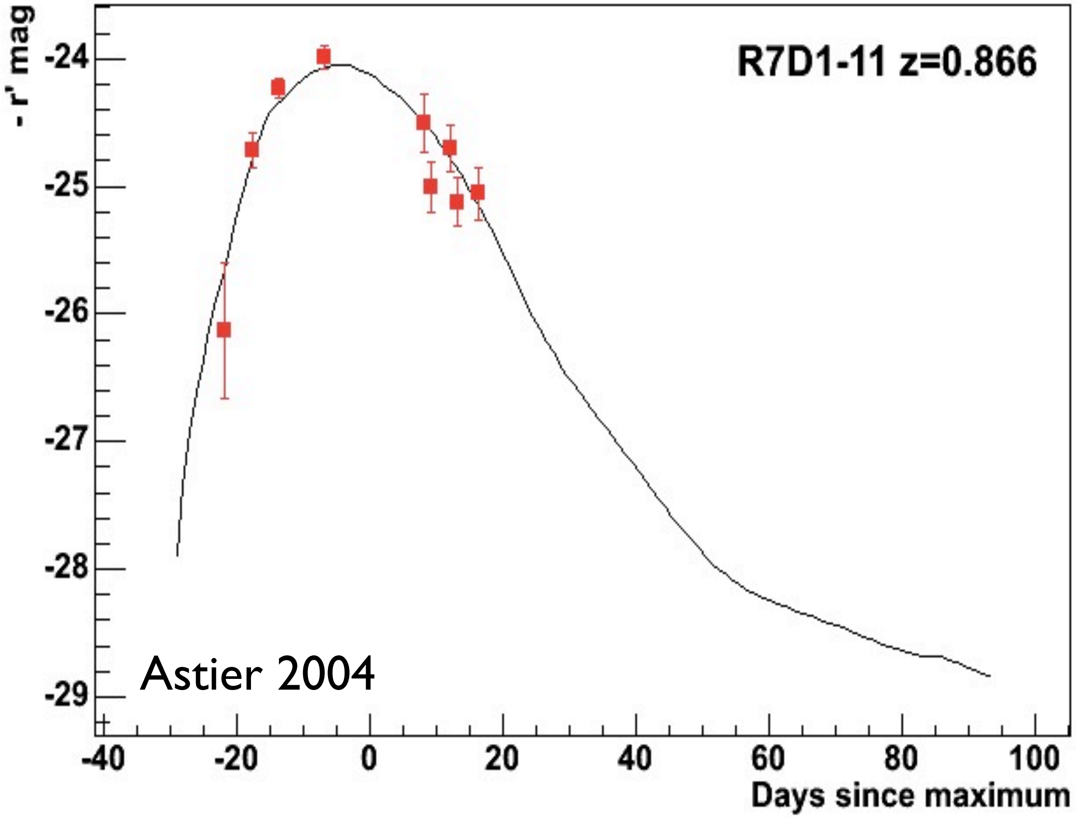
How to use SNe Ia data (SALT II) to do cosmological model selection within the Bayesian framework



Marisa Cristina March

University of Sussex

m.march@sussex.ac.uk

1102.3237v4 & 1207.3705v1

Bayes equation for SNe Ia

$$p(\theta|d,\mathcal{M}) = \frac{p(d|\theta,\mathcal{M})p(\theta|\mathcal{M})}{p(d|\mathcal{M})} \quad (1)$$

$$p(\theta|\underline{\hat{c}},\underline{\hat{x}}_1,\underline{\hat{m}}_B^*) = \frac{p(\underline{\hat{c}},\underline{\hat{x}}_1,\underline{\hat{m}}_B^*|\theta,\mathcal{M})p(\theta|\mathcal{M})}{p(\underline{\hat{c}},\underline{\hat{x}}_1,\underline{\hat{m}}_B^*|\mathcal{M})} \quad (2)$$

- data $d = \{\underline{\hat{c}},\underline{\hat{x}}_1,\underline{\hat{m}}_B^*\}$
- parameters $\theta = \{\Omega_m,\Omega_\Lambda,w,\alpha,\beta,M_0\}$
- Stretch and colour relation $\underline{\mu} = \underline{m}_B^* - M_0 + \alpha\underline{x}_1 - \beta\underline{c}$
- Question: What is $p(\underline{\hat{c}},\underline{\hat{x}}_1,\underline{\hat{m}}_B^*|\theta,\mathcal{M})$?

m.march@sussex.ac.uk
1102.3237v4 & 1207.3705v1

Standard (χ^2) method

As a distance estimator, we use:

$$\mu_B = m_B^* - M + \alpha(s-1) - \beta c$$

where m_B^*, s and c are derived from the fit to the light curves, and α, β and the absolute magnitude M are parameters which are fitted by minimizing the residuals in the Hubble diagram. The cosmological fit is actually performed by minimizing:

$$\chi^2 = \sum_{\text{objects}} \frac{(\mu_B - 5\log_{10}(d_L(\theta, z)/10\,\text{pc}))^2}{\sigma^2(\mu_B) + \sigma_{\text{int}}^2},$$

where θ stands for the cosmological parameters that define the fitted model (with the exception of H_0), d_L is the luminosity distance, and σ_{int} is the intrinsic dispersion of SN absolute magnitudes. We minimize with respect to θ, α, β and M.

Since d_L scales as $1/H_0$, only M depends on H_0. The definition of $\sigma^2(\mu_B)$, the measurement variance, requires some care. First, one has to account for the full covariance matrix of m_B^*, s and c from the light-curve fit. Second, $\sigma(\mu_B)$ depends on α and β; minimizing with respect to them introduces a bias towards increasing errors in order to decrease the χ^2, as originally noted in Tripp (1998). When minimizing, we therefore fix the values of α and β entering the uncertainty calculation and update them iteratively. $\sigma(\mu_B)$ also includes a peculiar velocity contribution of 300 km s^{-1}. σ_{int} is introduced to account for the "intrinsic dispersion" of SNe Ia. We perform a first fit with an initial value (typically 0.15 mag), and then calculate the σ_{int} required to obtain a reduced $\chi^2 = 1$. We then refit with this

- Above extract from Astier 2006
- $\sigma^2(\mu_B)$ is a funtion of α, β
- $p(\underline{\hat{c}}, \underline{\hat{x}}_1, \underline{\hat{m}}_B^*|\theta) \neq \chi^2$

m.march@sussex.ac.uk

What is $p(\underline{\hat{c}}, \underline{\hat{x}}_1, \underline{\hat{m}}^*_B|\theta)$?

$$
\begin{aligned}
p(\underline{\hat{c}}, \underline{\hat{x}}_1, \underline{\hat{m}}^*_B|\theta) &= \int \prod_i p(\hat{c}_i, \hat{x}_{1i}, \hat{m}^*_{Bi}|c_i, x_{1i}, m^*_{Bi}, \theta) \\
&\times p(m^*_{Bi}|c_i, x_{1i}, \theta) p(c_i, x_{1i}|\theta)\ \mathrm{dc_i}\ \mathrm{dx_{1i}}\ \mathrm{dm^*_{Bi}} \\
&= \int \prod_i (2(\pi\sigma^2_{m,i} + \sigma_\mu^{\text{int},2}))^{-\frac{1}{2}} \exp\left(-\frac{1}{2}\frac{(m^*_{Bi} - \hat{m}^*_{Bi})^2}{(\sigma^2_{m,i} + \sigma_\mu^{\text{int},2})}\right) \\
&\times (2\pi\sigma^2_{c,i})^{-\frac{1}{2}} \exp\left(-\frac{1}{2}\frac{(c_i - \hat{c}_i)^2}{(\sigma^2_{c,i})}\right) (2\pi\sigma^2_{x,i})^{-\frac{1}{2}} \exp\left(-\frac{1}{2}\frac{(x_{1i} - \hat{x}_{1i})^2}{(\sigma^2_{c,i})}\right) \\
&\times \delta(m^*_{Bi} - (\mu_i + M_0 - \alpha x_{1i} + \beta c_i)) \\
&\times p(c_i|\theta) \\
&\times p(x_{1i}|\theta)\ \mathrm{dc_i}\ \mathrm{dx_{1i}}\ \mathrm{dm^*_{Bi}}
\end{aligned}
$$

What do we choose for $p(x_{1i}|\theta)$ and $p(c_i|\theta)$?

m.march@sussex.ac.uk

1102.3237v4 & 1207.3705v1

Choice 1: 'Uninformative' prior

- Choose broad flat priors on latent c_i, x_{1i}
- $p(x_{1i}|\theta) = k_x$ and $p(c_i|\theta) = k_c$
- After $\mathrm{dm}^*_{\mathrm{Bi}}$ integral, we have:

$$p(\underline{\hat{c}}, \underline{\hat{x}}_1, \underline{\hat{m}}^*_B|\theta) = \int \prod_i (2\pi(\sigma^2_{m,i} + \sigma_\mu^{\mathsf{int},2}))^{-\frac{1}{2}} \exp\left(-\frac{1}{2}\frac{((\mu_i + M_0 - \alpha x_{1i} + \beta c_i) - \hat{m}^*_{Bi})^2}{(\sigma^2_{m,i} + \sigma_\mu^{\mathsf{int},2})}\right)$$
$$\times (2\pi\sigma^2_{c,i})^{-\frac{1}{2}} \exp\left(-\frac{1}{2}\frac{(c_i - \hat{c}_i)^2}{(\sigma^2_{c,i})}\right) (2\pi\sigma^2_{x,i})^{-\frac{1}{2}} \exp\left(-\frac{1}{2}\frac{(x_{1i} - \hat{x}_{1i})^2}{(\sigma^2_{x,i})}\right)$$
$$\times k_c \times k_x \; \mathrm{dc_i} \; \mathrm{dx_{1i}}$$

- Next do integrals over $\mathrm{dc_i}\ \ \mathrm{dx_{1i}}$ integral, to obtain......

University of Sussex

m.march@sussex.ac.uk
1102.3237v4 & 1207.3705v1

Choice 1: 'Uninformative' prior

- After $\mathrm{dc_i}\ \ \mathrm{dx_{1i}}$ integrals, we have:

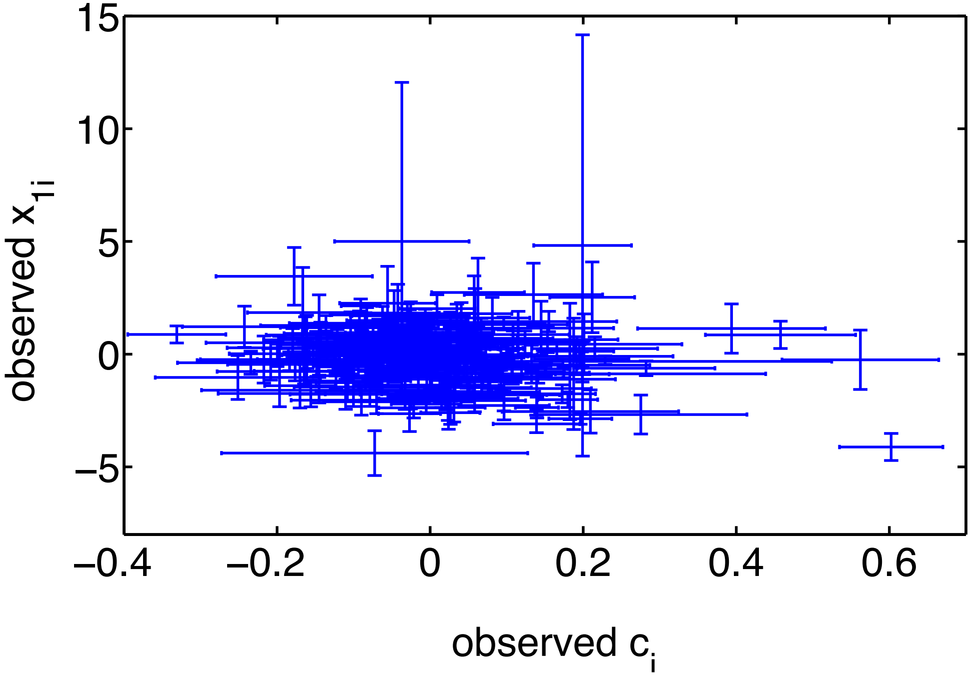
$$p(\underline{\hat{c}}, \underline{\hat{x}}_1, \underline{\hat{m}}_B^*|\theta) = \prod_i (2\pi(\sigma_{m,i}^2 + \sigma_\mu^{\text{int},2} + \alpha^2\sigma_{x,i}^2 + \beta^2\sigma_{c,i}^2))^{-\frac{1}{2}}$$
$$\times \exp\left(-\frac{1}{2}\frac{(\mu_i + M_0 - \alpha\hat{x}_{1i} + \beta\hat{c}_i - \hat{m}_{Bi}^*)^2}{(\sigma_{m,i}^2 + \sigma_\mu^{\text{int},2} + \alpha^2\sigma_{x,i}^2 + \beta^2\sigma_{c,i}^2)}\right)$$
$$\times k_c \times k_x$$

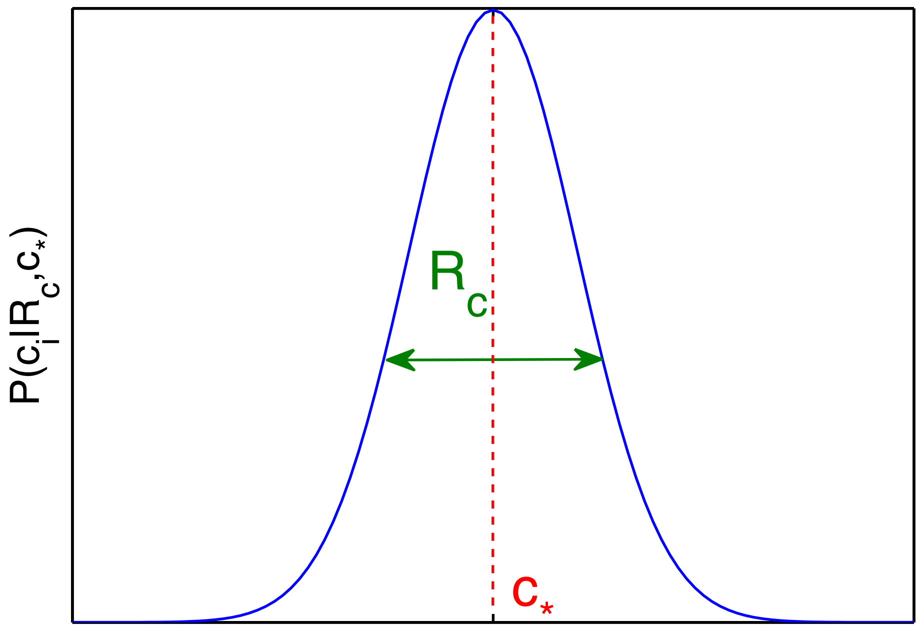
- Exponential term is identical to χ^2 of standard method
- Using this expression results in a biased estimator for α
- The method is correct, however, the choice of prior is incorrect.

University of Sussex

m.march@sussex.ac.uk

Which priors on the latent c_i, x_{1i}?





- Size of the error bars on $\hat{x}_{1i}$ and $\hat{c}_i$ is comparable with range of $\hat{x}_{1i}$ and $\hat{c}_i$.
- If not properly accounted for, this leads to a bias in the recovery of the SN parameters α and β.
- Solution is to put an informative prior the latent $x_1 i$ and c_i. [Gull, 1989].

m.march@sussex.ac.uk

1102.3237v4 & 1207.3705v1

Choice 2: ‘Informative’ prior

- Choose Gaussian priors on latent $p(c_i|c_\star, Rc)$ and $p(x_{1i}|x_\star, Rx)$
- After $\mathrm{dm}^*_{\mathrm{Bi}}$ integral, we have:

$$p(\underline{\hat{c}}, \underline{\hat{x}}_1, \underline{\hat{m}}^*_B|\theta) = \int \prod_i (2\pi(\sigma^2_{m,i} + \sigma_\mu^{\text{int},2}))^{-\frac{1}{2}} \exp\left(-\frac{1}{2}\frac{((\mu_i + M_0 - \alpha x_{1i} + \beta c_i) - \hat{m}^*_{Bi})^2}{(\sigma^2_{m,i} + \sigma_\mu^{\text{int},2})}\right)$$
$$\times (2\pi\sigma^2_{c,i})^{-\frac{1}{2}} \exp\left(-\frac{1}{2}\frac{(c_i - \hat{c}_i)^2}{(\sigma^2_{c,i})}\right)(2\pi\sigma^2_{x,i})^{-\frac{1}{2}} \exp\left(-\frac{1}{2}\frac{(x_{1i} - \hat{x}_{1i})^2}{(\sigma^2_{x,i})}\right)$$
$$\times (2\pi R_c^2)^{-\frac{1}{2}}) \exp\left(-\frac{1}{2}\frac{(c_i - c_\star)^2}{(R_c^2)}\right)$$
$$\times (2\pi R_x^2)^{-\frac{1}{2}}) \exp\left(-\frac{1}{2}\frac{(x_{1i} - x_\star)^2}{(R_x^2)}\right)$$
$$\times p(c_\star)p(x_\star)p(R_c)p(R_x)\ \mathrm{dc_i}\ \mathrm{dx_{1i}}\ \mathrm{dc_\star}\ \mathrm{dx_\star}\ \mathrm{dR_c}\ \mathrm{dR_x}$$

m.march@sussex.ac.uk
1102.3237v4 & 1207.3705v1

Choice 2: 'Informative' prior

- Integrate over latent parameters $\underline{c}, \underline{x}_1$ analytically.
- Integrate over population parameters $c_\star, x_\star$ analytically.

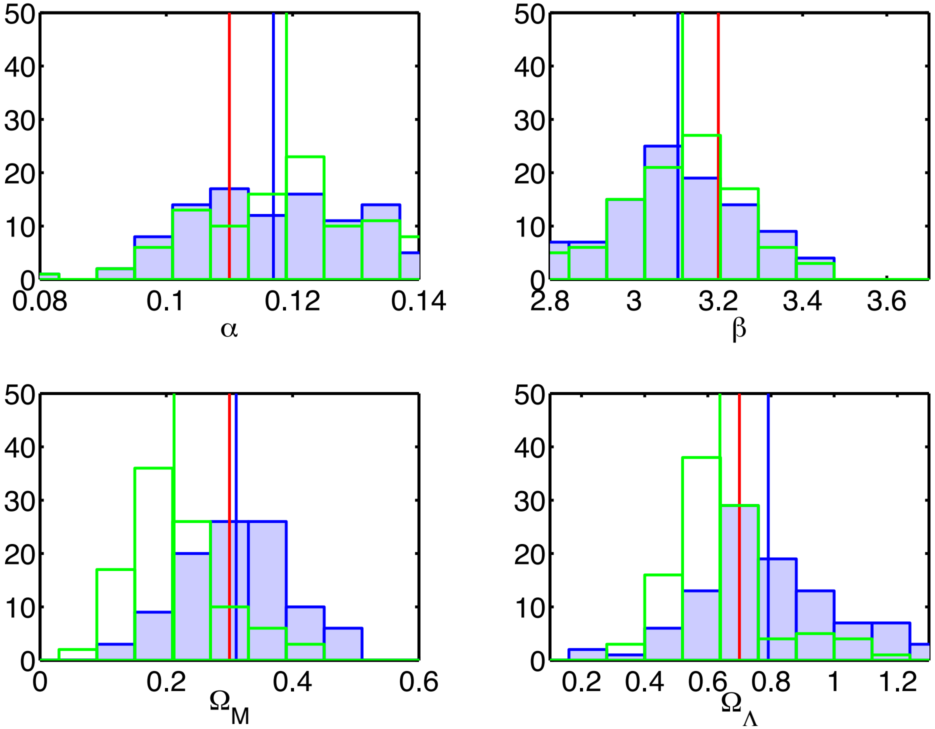
to obtain effective likelihood:

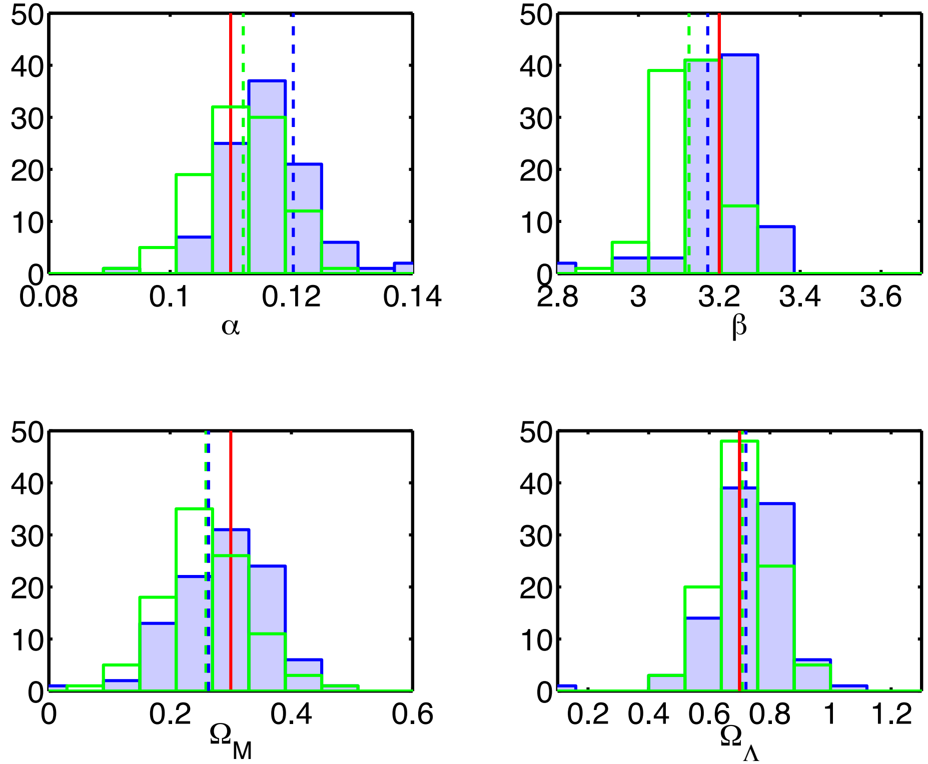
$$p(\underline{\hat{c}}, \underline{\hat{x}}_1, \underline{\hat{m}}_B^*|\Theta) = \int \mathsf{d}\log R_c \ \mathsf{d}\log R_x \ |2\pi\Sigma_C|^{-\frac{1}{2}}|2\pi\Sigma_P|^{-\frac{1}{2}}|2\pi\Sigma_A|^{\frac{1}{2}}|2\pi\Sigma_0|^{-\frac{1}{2}}|2\pi K|^{\frac{1}{2}}$$
$$\times \exp\left(-\frac{1}{2}[X_0^T\Sigma_C^{-1}X_0 - \Delta^T\Sigma_A\Delta - k_0^TK^{-1}k_0 + \underline{b}_m^T\Sigma_0^{-1}\underline{b}_m]\right)$$

- Integrate over population parameters R_c, R_x numerically.
- This choice of prior gives the correct likelihood for $p(\underline{\hat{c}}, \underline{\hat{x}}_1, \underline{\hat{m}}_B^*|\Theta)$ necessary for using SNe Ia to do cosmological model selection within the Bayesian framework.

m.march@sussex.ac.uk

Numerical trials with simulated data





- LH panel: SNLS3 only data; RH panel, 'cosmology sample'
- Blue: Bayesian method; Green, χ^2 method

- Histograms show estimators from 100 realizations of simulated data.

University of Sussex

m.march@sussex.ac.uk
1102.3237v4 & 1207.3705v1

Where the two methods meet

Bayesian Data Analysis: Straight-line fitting

Stephen F. Gull
Cavendish Laboratory,
Madingley Road,
Cambridge CB3 OHE, U.K.

Abstract

A Bayesian solution is presented to the problem of straight-line fitting when both variables x and y are subject to error. The solution, which is fully symmetric with respect to x and y, contains a very surprising feature: it requires a <u>informative</u> prior for the distribution of sample positions. An uninformative prior leads to a bias in the estimated slope.

Can such a simple problem really
you want is the answer I can recommend

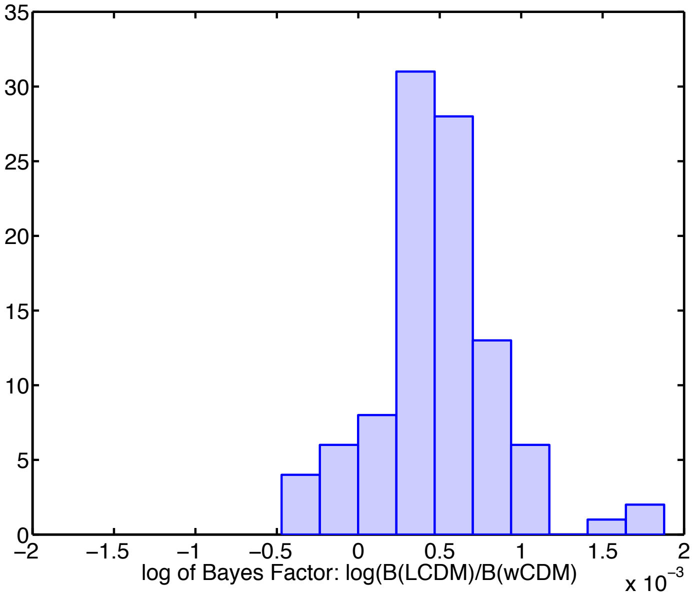
$$\min \frac{(a\, V_{xx} - 2\, V_{xy} + a^{-1}\, V_{yy})}{(a\, \sigma_x^2 + a^{-1}\, \sigma_y^2)} \;.$$

This is our answer with R --> ∞
will probably work. In can be derived
private communication, see also Ripl

- Above extract from Gull 1989. The χ^2 method is an approximation to the full Bayesian formalism, in some limit, to be determined.
- To do Bayesian model selection, must use full expression in order to obtain correct evidence calculation.

m.march@sussex.ac.uk

Numerical Trials: Bayes Factor



- Log of the ratio of Bayesian evidence between ΛCDM and flat wCDM, for 100 sets of simulated data.