COSMOLOGY: DATA ANALYSIS

We have a set of data $\vec{x}$; we will obtain how much H_0?

you have a model M, you have set of parameters $\vec{\theta}$

$\Downarrow$

You want to obtain the probability (of having) distribution of the parameter at given data

$$P(\vec{\theta}\,|\,\vec{x})$$

CONDITIONAL PROBABILITY

what is the mean value, what is its uncertainty

DATA + MODEL $\Rightarrow$ YOU CAN DO EXPERIMENT OVER AND OVER
FREQUENTIST APPROACH

COSMOLOGY ??? OBSERVATION LOOK UP $\rightarrow$ THEN WE HAVE A SET OF DATA AND GAME IS OVER

$P(\vec{x}\,|\,\vec{\theta})$: Prop of a data given for a parameter.

So we should find a way (statistics) to do the inverse $\Rightarrow$

Easy to compute LIKELYHOOD

POSTERIOR

$$P(\vec{\theta}|\vec{x}) = \frac{P(\vec{x}\,|\,\vec{\theta})}{\underbrace{P(\vec{x})}_{\text{Joint Probability}}} = \frac{P(\vec{x}\,|\,\vec{\theta})\,P(\vec{\theta})}{\underbrace{P(\vec{x})}_{\text{Evidence}}}$$

PRIOR

Evidence : quite useful for model selection

$$-\ln \underbrace{P(\vec{x}\,|\,\theta)}_{\text{Prob of your data for a given parameter}} \propto -\frac{1}{2}(\vec{x} - x_{TH}(\theta))^T C^{-1} (\vec{x} - \vec{x}_{TH}(\vec{\theta}))$$

What you get

what do you expect with a given parameter

COVARIANCE MATRIX
YOU HAVE MEASUREMENT ERROR

Gaussian distribution [Assume

$$\int P(\vec{\theta}\,|\,\vec{x})\,d\theta = 1 \Rightarrow P(x) = \int P(x\,|\,\theta)\,P(\theta)\,d\theta$$

NORMALIZATION

$P(\theta)$: PRIOR KNOWLEDGE ON THE PARAMETER

$$P(H_0) \propto \text{EXP}\left[\frac{(H_0 - \hat{H}_0)^2}{\sigma_{H_0}^2}\right]$$

supernova measurement mean

supernova measurement's error

$\hat{H}_0 = 70 \pm 1$

$$\propto \exp\left(\frac{(H_0 - 70)^2}{1^2}\right)$$

H_0 is about 50 and

come on!

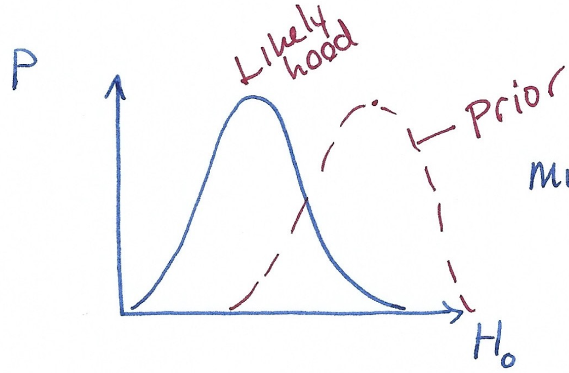
No $H_0 \sim 70$

Supernova

$$\underbrace{P(H_0\,|\,\vec{x})}_{} = \frac{P(\vec{x}\,|\,H_0)\,P(H_0)}{P(\vec{x})}$$

Ohh
life is a prettier easy

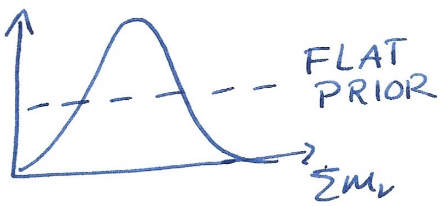
Thanks God I have prior!!!



Multiplying the two

You are measuring neutrino mass from CMB,

$\sum m_\nu \geq 0$



$P(\sum m_\nu) \propto \theta(\sum m_\nu)$

all values are equally likely

ESTIMATE (?) – before Measurement you need to find estimator of the parameter.

$\mathcal{L}$ gaussian and prior

$\theta_{\text{max likelyhood}}$: $\max \mathcal{L}(x|\theta)$ PEAK

$$\hat{\theta} = \int d\vec{\theta}\, \vec{\theta}\, P(\vec{\theta}|\vec{x})$$ = MEAN EXPECTED VALUE

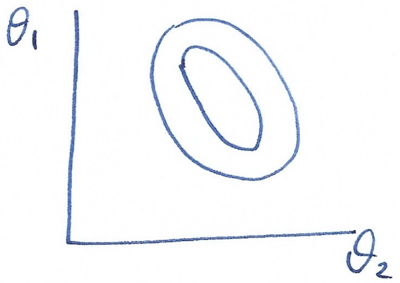
Estimator

MARGINALIZATION : you want to find Prob. distribution of a single Parameter

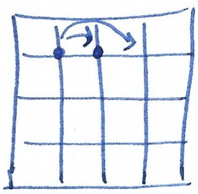
$$P(\theta_1|\vec{x}) = \int \underbrace{d\theta_2 \dots d\theta_N}_{\text{integrated over all remaining parameter.}} P(\theta|\vec{x})$$

$$\int_{\theta_{min}}^{\theta_{max}} d\theta_1 \, P(\theta_1 | x) = 0.683$$

TWO PARAMETERS → CONTOURS



IF YOU HAVE ONLY TWO PARAMETERS (θ_1 and θ_2)



$P(\theta|x)$ HOW DO WE FIND?

for θ_1 have a value and θ_2 may have a value and calculate one by one

6 parameters of ΛCDM

MONTE CARLO APPROACH } you want to move in your parameter space.

METROPOLIS–HASTINGS ALGORITHM TO MOVE] A particular one

1) chose $\vec{\theta}_0 \longrightarrow P(\vec{\theta}_0 | \vec{x})$ compute

2) move from θ_0 to $\quad N(\vec{\theta}) \propto \mathrm{Exp}\left[-\frac{1}{2}\frac{(\vec{\theta}-\vec{\theta}_0)^2}{\sigma_{\theta_0}^2}\right] \longrightarrow \vec{\theta}_1 = P(\vec{\theta}_1|\vec{x})$

3) $a = \dfrac{P(\theta_1|x)}{P(\theta_0|x)}$ what does it say? >1 $P(\theta_1|x) > P(\theta_0|x)$ you are going to highest probability (accept it)

$$\hat{\theta}_c = \frac{1}{N} \sum_{i=1} \theta_i^c \qquad \text{mean of c-chain}$$

$$\hat{\theta} = \frac{1}{N} \frac{1}{M} \sum_{i=1}^{N} \sum_{c=1}^{M} \theta_i^c \qquad \text{mean of all chains}$$

$$W = \frac{1}{M} \frac{1}{N-1} \sum_{i=1}^{N} \sum_{c=1}^{M} \left(\theta_i^c - \hat{\theta}^c \right)^2 \qquad \text{average of each chain variance}$$

$$\frac{B}{N} = \frac{1}{M-1} \sum_{c=1}^{M} \left(\hat{\theta}^c - \hat{\theta} \right)^2 \qquad \text{chain to chain variance.}$$

The unbiased estimator of the variance of θ is

$$V = \frac{N-1}{N} W + \frac{B}{N} \left(1 + \frac{1}{M} \right)$$

$$\frac{V}{W} = \frac{N-1}{N} + \frac{B}{N} W \left[1 + \frac{1}{M} \right]$$

$$V \simeq W$$

unbiased estimator

average of each chain variance

when the chain is converged $V \simeq W$

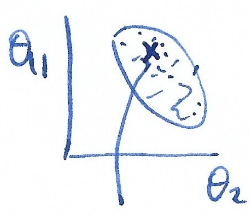
$$R = \frac{V}{W} \longrightarrow 1$$

$R = \frac{V}{W} \geq 1$ } chain should progress

till $R-1$ approaches ~~[illegible]~~ zero.

OK CHAIN IS CONVERGED

STOP I GOT.



mean

Remove the column

it mean it's marginalized.

$$\text{ratio} = \frac{P(\theta_1|x)}{P(\theta_0|x)}$$

- ratio > 1 accept $\theta_0^1 \dots \theta_0^2 \dots$ / $\theta_1^1 \dots \theta_1^2 \dots$ chain
- ratio < 1 $R \in [0,1]$ in this case you are defining R (random)
  - $R < a$ accept.
  - $R > a$ reject the point. kill this process and start from $\vec{\theta}_0$ again.

IF YOU LET THIS CHAIN GOING ON LONG ENOUGH

$$\rho_{\theta_i \in \Omega} \propto \int_\Omega d\vec{\theta}\, P(\vec{\theta}|\vec{x})$$

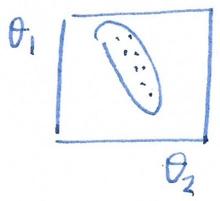
fraction of accepted points in a given region $\Longleftrightarrow$ Prob of finding parameter values in the region.

WHAT DOES THE LONG ENOUGH MEAN?

TWO, FOUR DIFFERENT CHAINS,

within a chain
covariance of a chain
covariance chain to chain.
chains converge together.

CONFIDENCE REGIONS



Marginalize

$\theta_0^1 \dots \theta_0^2 \dots$
$\theta_1^1 \dots \theta_1^2 \dots$

Remove this column

Unfortunately, this is not the end of the story!

Chains are constructed Metropolis Hasting algorithm.

To assess that you reached convergence = THE TOOL IS GELMAN-RUBIN STATISTICS

M chains

N accepted point each chain.

M counted by c

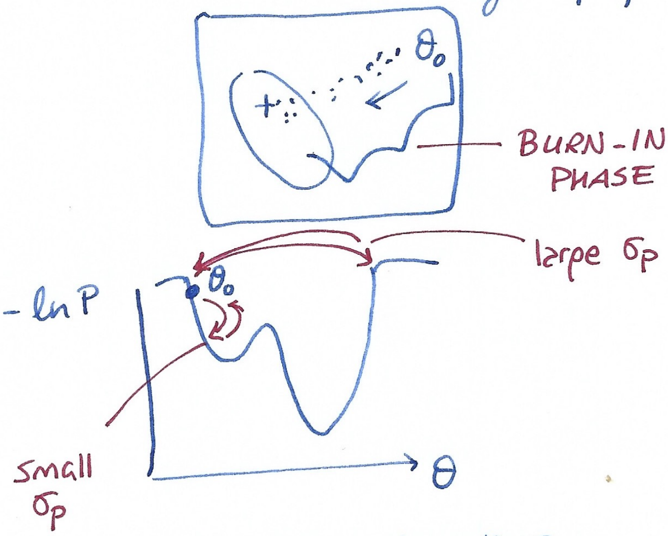
i th point in the c-th chain

N counted by i

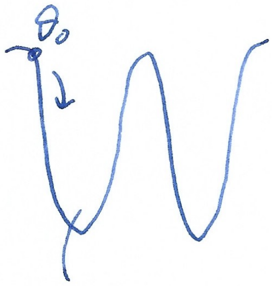
$$\hat{\theta}_c = \frac{1}{N}\sum_{i=1}^{N} \theta_i^c \quad \text{mean of c-chain}$$

you have to decide where you start θ_0,

// your proposal distribution σ_p



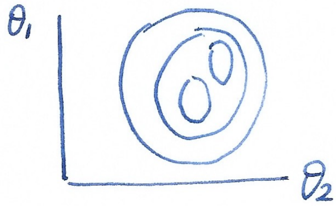
the reason that we have more chains



you explore here and you cannot explore the second.

in MH algorithms

to overcome this issue;
NESTED SAMPLING
IS PROPOSED



MODEL SELECTION

$$P(M|\vec{X}) \propto \underbrace{P(\vec{X}|M)}_{\int d\theta\, P(\vec{x}|\vec{\theta})\, P(\vec{\theta})} P(M)$$

if you select one, you need at least two.

COMPARISON = RATIO

$$\frac{P(M_2|\vec{x})}{P(M_1|\vec{x})} = \frac{P(M_2)}{P(M_1)} \frac{\int d\theta\, \cancel{\ldots}\, P(\vec{x}|\theta, M_2)\, P(\theta|M_2)}{\int d\theta\, \cancel{\ldots}\, P(\vec{x}|\theta, M_1)\, P(\theta|M_1)}$$

BIAS RATIO = $B = \dfrac{\int d\theta\, \mathcal{L}(\vec{x}|\theta, M_2)\, P(\theta, M_2)}{\int d\theta\, \mathcal{L}(\vec{x}|\theta, M_1)\, P(\theta, M_1)}$

If we do not have no a priori preference for one model over the other $P(M_2) \simeq P(M_1)$

we can get

B BAYES FACTOR $\longrightarrow$ goodness of fit (χ^2) & complexity of the model

PEAK OF POSTERIOR

let us remember Neumann's words] χ^2 fit is not enough

More complilated models (more parameters) would be always favoured by a goodness of fit analysis, as they are able to accommodate the observations.

Yet, Bayes factor introduces a penalty for more complex models, giving weight to the enlarged prior volume necessary to fit the data.

$\theta_1^{i \to N}$

$$P(\theta_1 | M_1) = \left[\Delta\theta_1^1 \dots \Delta\theta_1^N \right]^{-1}$$

substituting to B

$p = N - N'$

M_1 has p parameter that are not in M_2

$$B = \frac{\int d\theta \, P(\vec{x} | \theta, M_2)}{\int d\theta \, P(\vec{x} | \theta, M_1)} \, \frac{\Delta\theta_1^1 \dots \Delta\theta_1^N}{\Delta\theta_2^1 \dots \Delta\theta_2^{N'}}$$

M_2 ΛCDM model $\quad H_0, \Omega_b h^2, \Omega_m$

M_1 wCDM $\quad H_0, \Omega_b h^2, \Omega_m + w$

$$B = \frac{\int d\theta \, P(\vec{x} | \theta_1, M_2)}{\int d\theta \, P(\vec{x} | \theta, M_1)} \, \Delta\theta_1^{N'+1} \dots \Delta\theta_1^{N'+p}$$

M_1 may provide better fit due to extra parameter;

$B < 1$ ($\swarrow$)

pushing towards a preference for the simpler model M_2 ΛCDM

Penalize the model with more parameters by increasing the value of B

Wrong choice of parameter and two datasets are disagreement MIGHT APPEAR AS CONCORDANT.

Existence of tensions between datasets

$$H_0^{CMB} = 67.35 \pm 0.54 \text{ km/s Mpc}$$
$$H_0^{SN} = 73.04 \pm 1.04 \text{ km/s Mpc}$$

measurements are distant from each other by around 5σ.

$$\frac{P(I_0 | D_1 \cup D_2, M)}{P(I_1 | D_1 \cup D_2, M)} = C = \text{Posterior ratio}$$

$$= \frac{P(D_1 \cup D_2 | I_0, M)}{P(D_1 \cup D_2 | I_1, M)} \cancel{\frac{P(I_0)}{P(I_1)}}$$

1

assuming.

$$= \frac{P(D_1 \cup D_2 | I_0, M)}{P(D_1 | M) P(D_2 | M)}$$

gaussian.

2 variables θ_1, θ_2 with P_1 and P_2;

Two experiments.

$$\underbrace{P(\theta_1 - \theta_2)}_{\text{probability of difference in measurements.}} = \int P_1(\tilde{\theta}) P_2(\tilde{\theta} - \Delta\theta)$$

tension $\Rightarrow$ $P(\Delta\theta)$ peak away from Zero.

OBSERVE A SHIFT $T^{obs} = \dfrac{|\theta(D_1) - \theta(D_2)|}{\sqrt{\sigma_\theta^2(D_1) + \sigma_\theta^2(D_2)}}$

$$P(T > T_{obs}) \propto \mathrm{Erf}\left(\frac{T_{obs}}{\sqrt{2}}\right)$$

useful method but it relies on a priori choice of parameter. so Be careful