BAYESIAN PARAMETER INFERENCE

DATA: $d \rightarrow \underline{d}$

PARAMETERS: $\theta \rightarrow \underline{\theta}$

MODEL: M

POSTERIOR DISTRIBUTION: $\Pr(\underline{\theta} | \underline{d}, M)$

LIKELIHOODS

$$= \frac{\Pr(\underline{\theta} | M) \Pr(\underline{d} | \underline{\theta}, M)}{\Pr(\underline{d} | M)} \quad \text{(BAYES'S THEOREM)}$$

$$= \frac{\Pr(\underline{\theta} | M) \Pr(\underline{d} | \underline{\theta}, M)}{\int d\underline{\theta}' \Pr(\underline{d} | \underline{\theta}', M) \Pr(\underline{\theta}' | M)} \quad \text{(LAW OF TOTAL PROBABILITY)}$$

$$\propto \Pr(\underline{\theta} | M) \Pr(\underline{d} | \underline{\theta}, M)$$

EXAMPLES

1. MONTY HALL PROBLEM (DISCRETE PARAMETER)

DATA: $h \in \{1, 2, 3\}$, $g \in \{1, 2, 3\}$

PARAMETERS: $p \in \{1, 2, 3\}$

MODEL: $M =$ "host follows rules"

$$\Pr(p | h, g, M) = \frac{1}{3} \delta_{p,g} + \frac{2}{3} (1 - \delta_{p,g})(1 - \delta_{p,h})$$

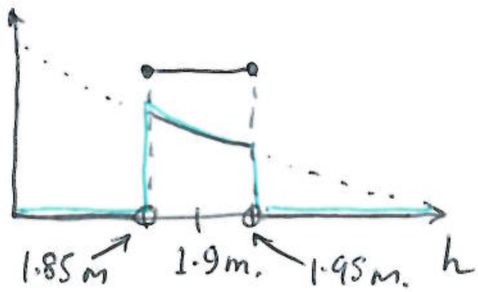
2. A PERSON'S HEIGHT (CONTINUOUS PARAMETER)

DATA: $\hat{h} = 1.9\,\text{m}$

PARAMETER: $h = ?$

MODEL: $M =$ "person is human"

PRIOR: $\Pr(h | M)$



3. COSMOLOGICAL PARAMETERS (MULTI-VARIATE)

DATA: $\mu_1\ \mu_2\ \mu_3 \ldots\ \{\mu_i\}\ \underline{\mu}\ ,\ z_1\ z_2\ z_3\ \{z_i\}\ \underline{z}\quad N_{SN}$

$\underline{d}$

PARAMETERS: $\underbrace{H_0\ \ \Omega_m}_{\underline{\theta}}$

MODEL: $M = \Lambda CDM$

$\hookrightarrow \mu_{th}(z)$

$\hookrightarrow \mu_{th}(z, H_0, \Omega_m, M)$

$Pr(\underline{\theta} \mid \underline{d}, M)$

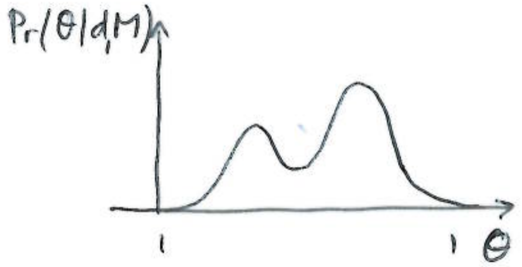
$\hookrightarrow Pr(H_0, \Omega_m \mid \{\mu_i, z_i\}, \Lambda CDM)$

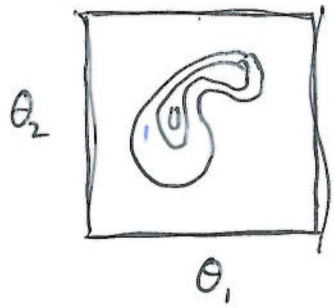
SUMMARY STATISTICS / GRAPHICS

$$\Pr(\underline{\theta} \mid \underline{d}, M)$$

MEAN: $\hat{\underline{\theta}} = \int d\underline{\theta}' \, \underline{\theta}' \, \Pr(\underline{\theta}' \mid \underline{d}, M)$

VARIANCE: $\underline{\underline{\sigma}}^2 = \int d\underline{\underline{\theta}}' \, (\underline{\underline{\theta}}' - \hat{\underline{\underline{\theta}}})^2 \, \Pr(\underline{\underline{\theta}}' \mid d, M)$





SAMPLING

$$\boxed{\underline{\theta}_i \sim \Pr(\underline{\theta} \mid \underline{d}, M)} \qquad i \in \{1, 2, \ldots N_{smp}\}$$

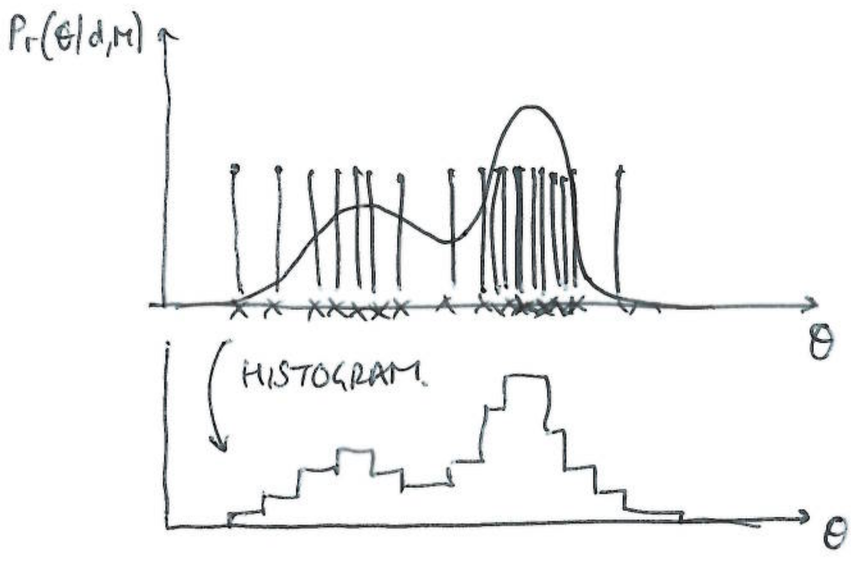
$$\{\underline{\theta}_i\}$$

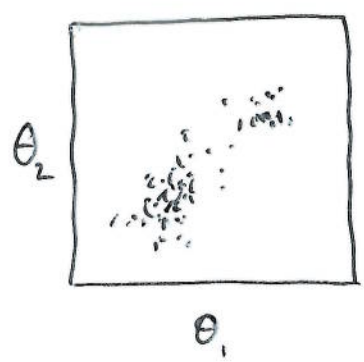
$$I = \int d\underline{\theta}' \, f(\underline{\theta}') \, \Pr(\underline{\theta}' \mid \underline{d}, M)$$

$$\hat{I} = \frac{1}{N_{smp}} \sum_{i=1}^{N_{smp}} f(\underline{\theta}_i) \qquad \text{if} \quad \underline{\theta}_i \sim \Pr(\underline{\theta} \mid \underline{d}, M)$$

$$\Rightarrow \quad \hat{\underline{\theta}} = \frac{1}{N_{smp}} \sum_{i=1}^{N_{smp}} \underline{\theta}_i$$

$$\tilde{\Pr}(\underline{\theta} \mid \underline{d}, M) = \frac{1}{N_{smp}} \sum_{i=1}^{N_{smp}} \delta_D(\underline{\theta} - \underline{\theta}_i)$$





MONTE CARLO SAMPLING TECHNIQUES

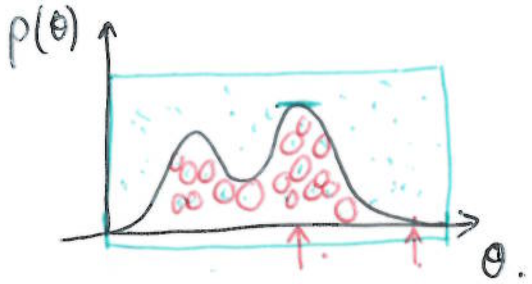
$$\Pr(\underline{\theta} \mid \underline{d}, M) \longrightarrow \rho(\underline{\theta}) \qquad \int \rho(\underline{\theta}')\, d\underline{\theta}' = 1.$$

$$\rho(\underline{\theta}) \geq 0 \quad \text{FOR ALL } \underline{\theta}$$

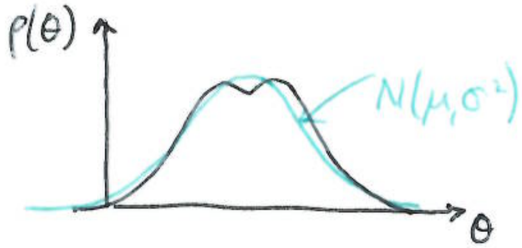
1. STANDARD DISTRIBUTIONS.

$$\rho(\theta) = N(\mu, \sigma^2) \qquad \text{(NORMAL)}$$

2. REJECTION SAMPLING



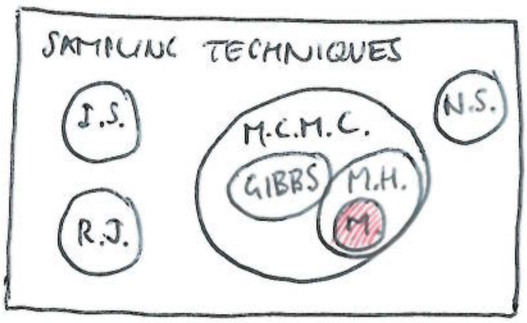
3. IMPORTANCE SAMPLING.



$$\theta_i \sim N(\mu, \sigma^2)$$

$$W_i = \frac{\rho(\theta_i)}{N(\theta_i; \mu, \sigma^2)}$$

4. MARKOV CHAIN MONTE CARLO. (MCMC)



METROPOLIS ALGORITHM.

$$\Pr(\underline{\theta} \mid \underline{d}, M) \longrightarrow \rho(\underline{\theta}) \longrightarrow \boxed{\rho'(\underline{\theta})} = C\,\rho(\underline{\theta})$$

$$\hookrightarrow \rho'(\underline{\theta}) = \Pr(\underline{\theta} \mid M)\,\Pr(\underline{d} \mid \underline{\theta}, M)$$

0: $\underline{\theta}_{start} \rightarrow \underline{\theta}_i$

1: $\underline{\theta}_{trial} \sim \underbrace{N(\underline{\theta}_i, \underline{\underline{\Sigma}})}_{\text{proposal distribution}}$ $\quad \underline{\underline{\Sigma}} = diag(\sigma^2, \sigma^2, \ldots \sigma^2) = \begin{bmatrix} \sigma^2 & & 0 \\ & \sigma^2 & \\ 0 & & \sigma^2 \end{bmatrix}$

2: ACCEPT WITH PROBABILITY:

$$P_{accept} \equiv \min\left(1, \frac{\rho'(\theta_{trial})}{\rho'(\theta_i)}\right) = \min\left(1, e^{\ln(\rho'(\theta_{trial})) - \ln(\rho'(\theta_i))}\right)$$

IF accept

$\underline{\theta}_{i+1} = \underline{\theta}_{trial}$ $\quad \leftarrow$ DETAILED BALANCE

ELSE

$\underline{\theta}_{i+1} = \underline{\theta}_i$ $\quad$ (REPEAT SAME VALUE)

GO TO STEP 1.

$$\frac{\rho(\theta_{trial})}{\rho(\theta_2)} \simeq \frac{1}{2}$$

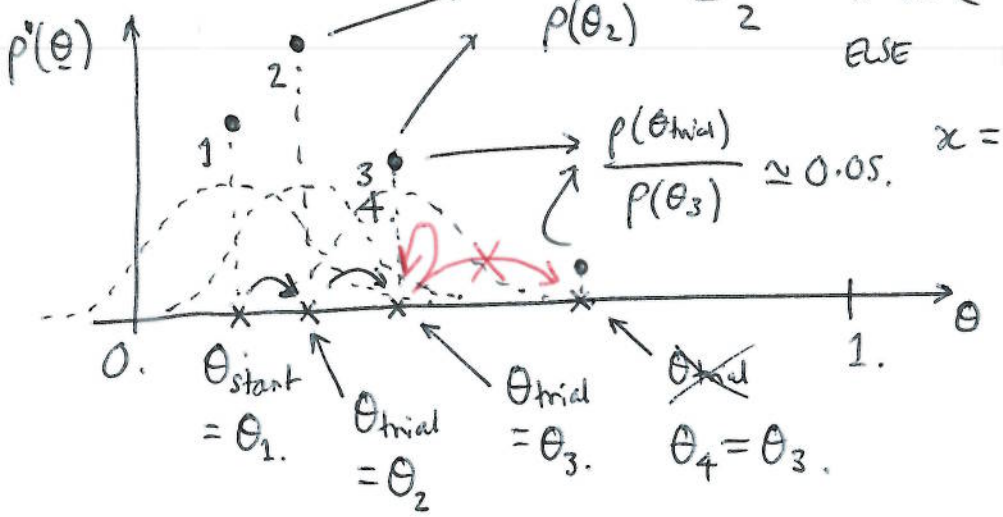
$x \sim U(0,1) \longrightarrow 0.14.$

IF $x \le 1/2$ accept

ELSE reject

$$\frac{\rho(\theta_{trial})}{\rho(\theta_3)} \simeq 0.05.$$

$x = 0.059$



$\theta_1 = 0.15$

$\theta_2 = 0.28$

$\theta_3 = 0.41$

$\theta_4 = 0.41$

$\theta_5 = 0.41$

LOG DENSITY

$$p_{accept} = \min\left[1, \frac{\rho'(\underline{\theta}_{trial})}{\rho'(\underline{\theta}_i)}\right]$$

$$= \min\left[1, \exp\left[\underbrace{\ln\left[\rho'(\underline{\theta}_{trial})\right]} - \underbrace{\ln\left[\rho'(\underline{\theta}_i)\right]}\right]\right]$$

$$L \propto \exp\Big[\underbrace{-\frac{1}{2}\sum_{i=1}\sum_{j=1}\ldots}_{\substack{\text{log-likelihood} \\ \ell(\underline{\theta})}}\Big]$$

$$p_{accept} = \min\left[1, \exp\left[\ell(\underline{\theta}_{trial}) - \ell(\underline{\theta}_i)\right]\right]$$

MARGINALISATION

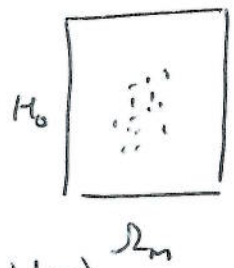
$$Pr(\underline{\theta} \mid \underline{d}, M) \qquad \theta_1 = H_0 \quad \theta_2 = \Omega_m$$

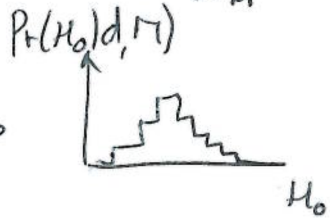
$$Pr(H_0, \Omega_m \mid \underline{d}, M)$$

$$Pr(H_0 \mid \underline{d}, M) = \int d\Omega_m \, Pr(H_0, \Omega_m \mid \underline{d}, M)$$

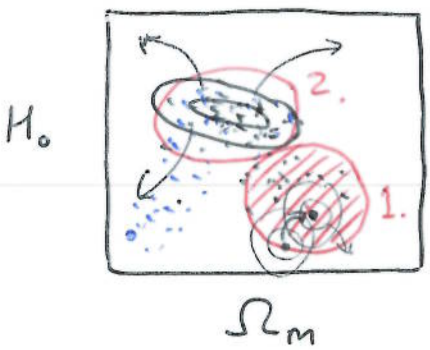
POTENTIALLY DIFFICULT

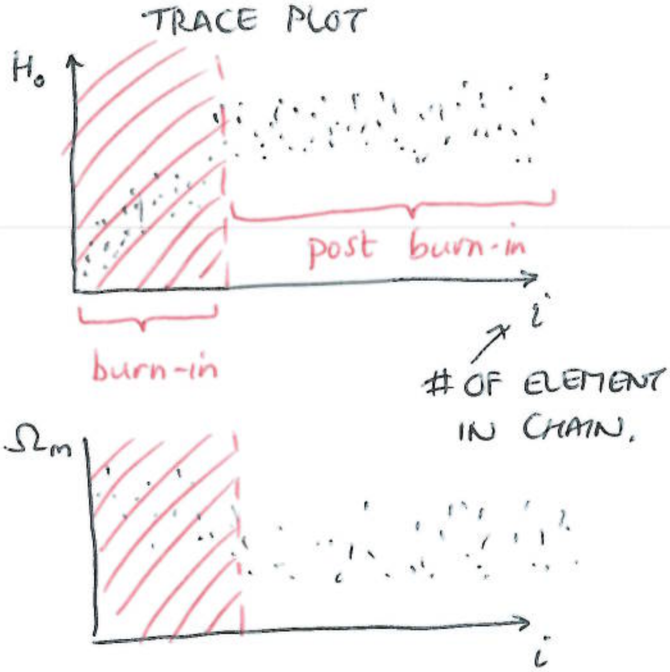
H_0	$\cancel{\Omega_m}$
71	$\cancel{0.3}$
72	$\cancel{0.4}$
72	$\cancel{0.4}$
72	$\cancel{0.4}$
73	$\cancel{0.5}$





BURN-IN



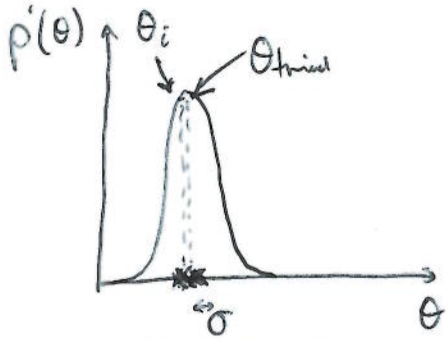


1. FIND ρ'_{max}
2. FIND FIRST SAMPLE WITH $\rho_i \geq 0.1 \, \rho'_{max}$
3. KEEP THE REST.

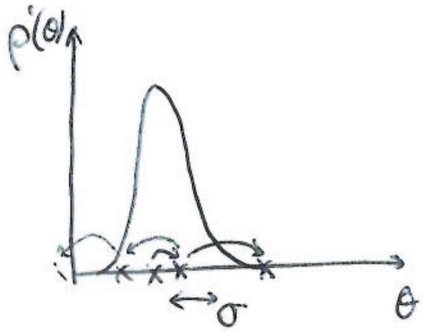
JUMP SIZE

1. $\theta_{trial} \sim N(\theta_i, \sigma^2)$ (ONE-DIMENSIONAL)

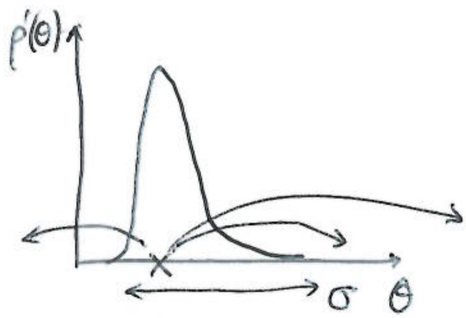
WHAT VALUE TO USE?



σ TOO SMALL



σ JUST RIGHT.



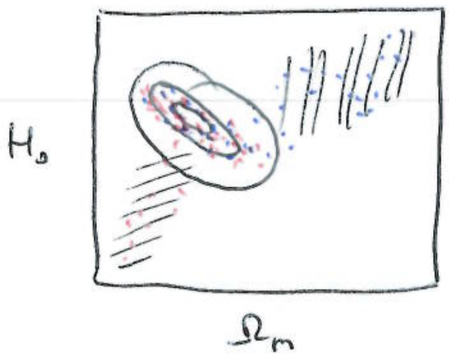
σ TOO LARGE

$$p_{accept} \simeq \min\left(1, \frac{\rho(\theta_{trial})}{\rho(\theta_i)}\right)$$

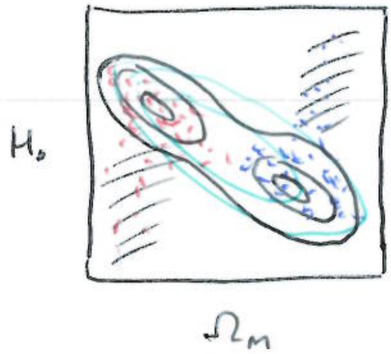
ACCEPTANCE RATIO
$\equiv$ FRACTION OF TIMES THAT θ_{trial} ACCEPTED

$\longrightarrow 1.$ $\qquad \sim \frac{1}{3}.$ $\qquad \longrightarrow 0.$

CONVERGENCE



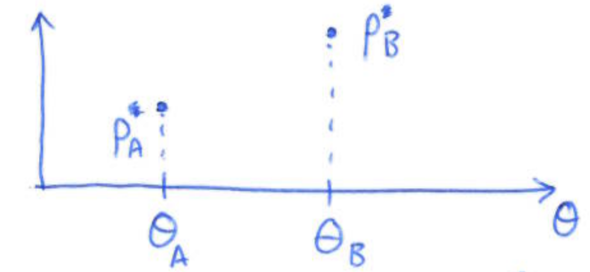
✓ CONVERGED



X NOT CONVERGED.

$\hat{R}$ STATISTIC (GELMAN & RUBIN 1992)

DETAILED BALANCE



$$\Pr(\theta_i = \theta_A) = \frac{p_A^*}{p_A + p_B} \qquad \Pr(\theta_i = \theta_B) = \frac{p_B}{p_A + p_B}.$$

$$\shortparallel$$

$$\Pr(\theta_{i+1} = \theta_A) = \frac{p_A}{p_A + p_B}.$$

$$= \underbrace{\Pr(\theta_i = \theta_A)} \Pr(A \to A) + \underbrace{\Pr(\theta_i = \theta_B)} \Pr(B \to A)$$

$$\frac{p_A}{\cancel{p_A + p_B}.} = \frac{p_A}{\cancel{p_A + p_B}} \left[1 - \Pr(A \to B)\right] + \frac{p_B}{\cancel{p_A + p_B}} \Pr(B \to A)$$

$$\cancel{p_A} = \cancel{p_A} - p_A \Pr(A \to B) + p_B \Pr(B \to A)$$

$$\boxed{\frac{\Pr(A \to B)}{\Pr(B \to A)} = \frac{p_B}{p_A}} \qquad p_B > p_A.$$

METROPOLIS: $\Pr(A \to B) = 1.$

✓. $\Pr(B \to A) = p_A / p_B$ $\Rightarrow$ $\dfrac{\Pr(A \to B)}{\Pr(B \to A)} = \dfrac{1}{p_A / p_B} = \dfrac{p_B}{p_A}.$

DETAILED BALANCE SATISFIED.

~~CRAP METROPOLIS~~ $\cancel{\Pr(A \to B) = 1/2}.$

~~(NOT TO BE USED)~~: $\cancel{\Pr(B \to A) = \frac{1}{2} \cdot \frac{p_A}{p_B}}.$

~~DISASTROUS METROPOLIS~~: $\cancel{\Pr(A \to B) = 0}$

$\cancel{\Pr(B \to A) = 0}.$