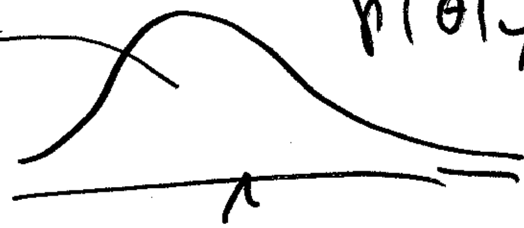
Monte Carlo (MC) data set

(8 Feb 13 extra notes (part 2)) ①

$p(\theta \mid y)$



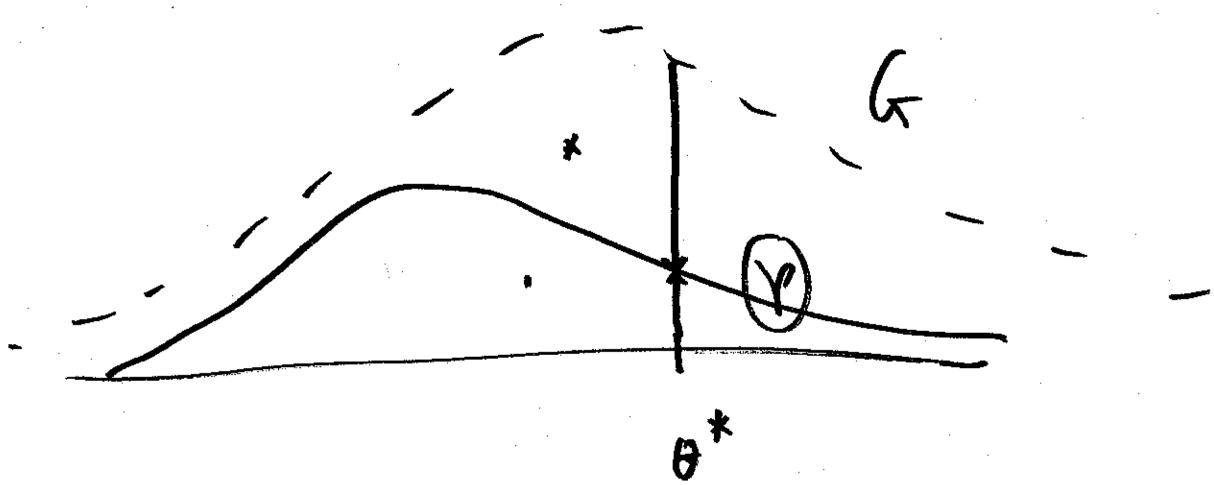
$\mu = E(\theta \mid y)$

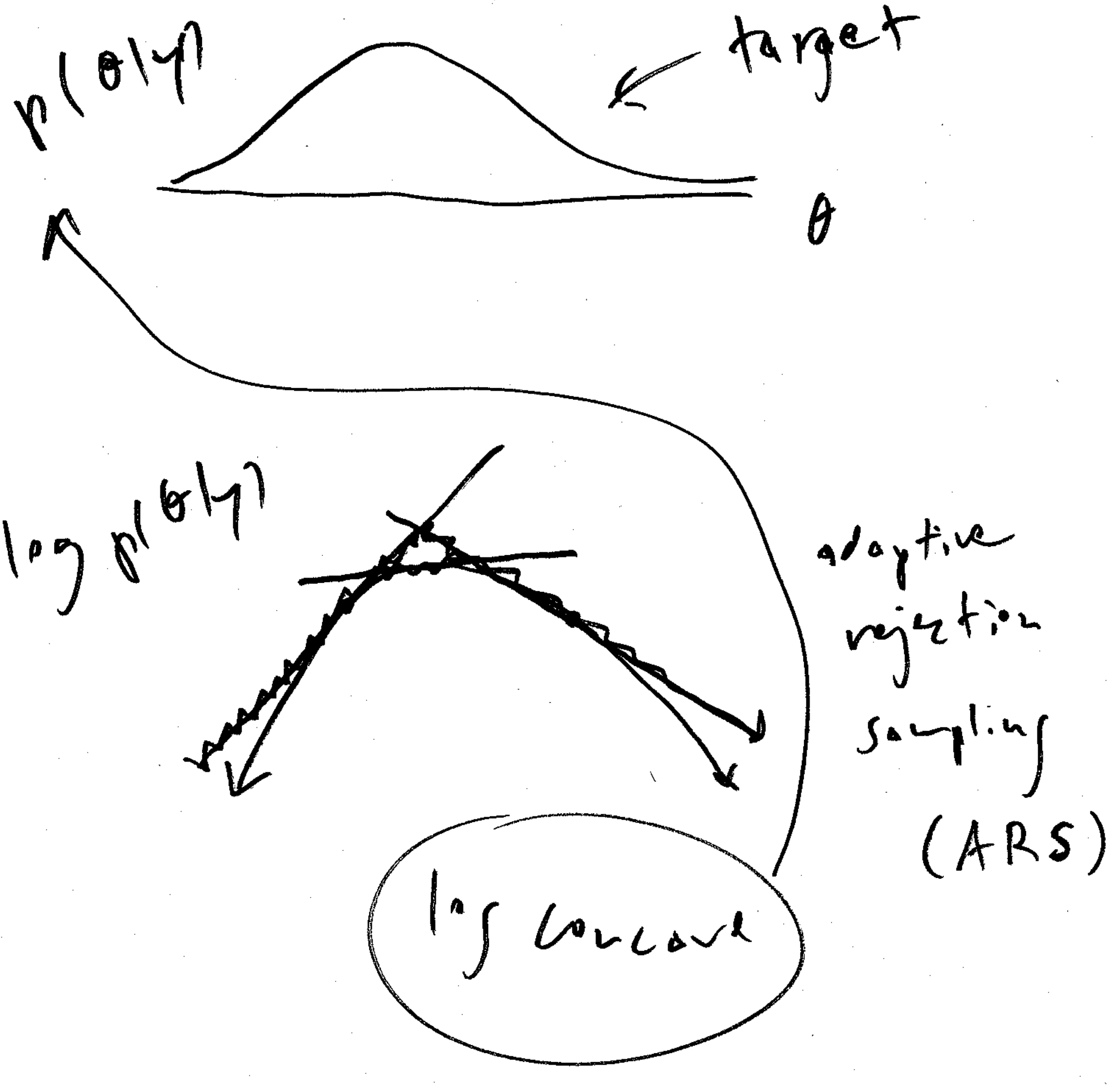
IID

$$\begin{bmatrix} \theta_1^* \\ \vdots \\ \theta_m^* \end{bmatrix}$$

(identically dist.)

mean $\bar{\theta}^*$ ← our MC est of μ

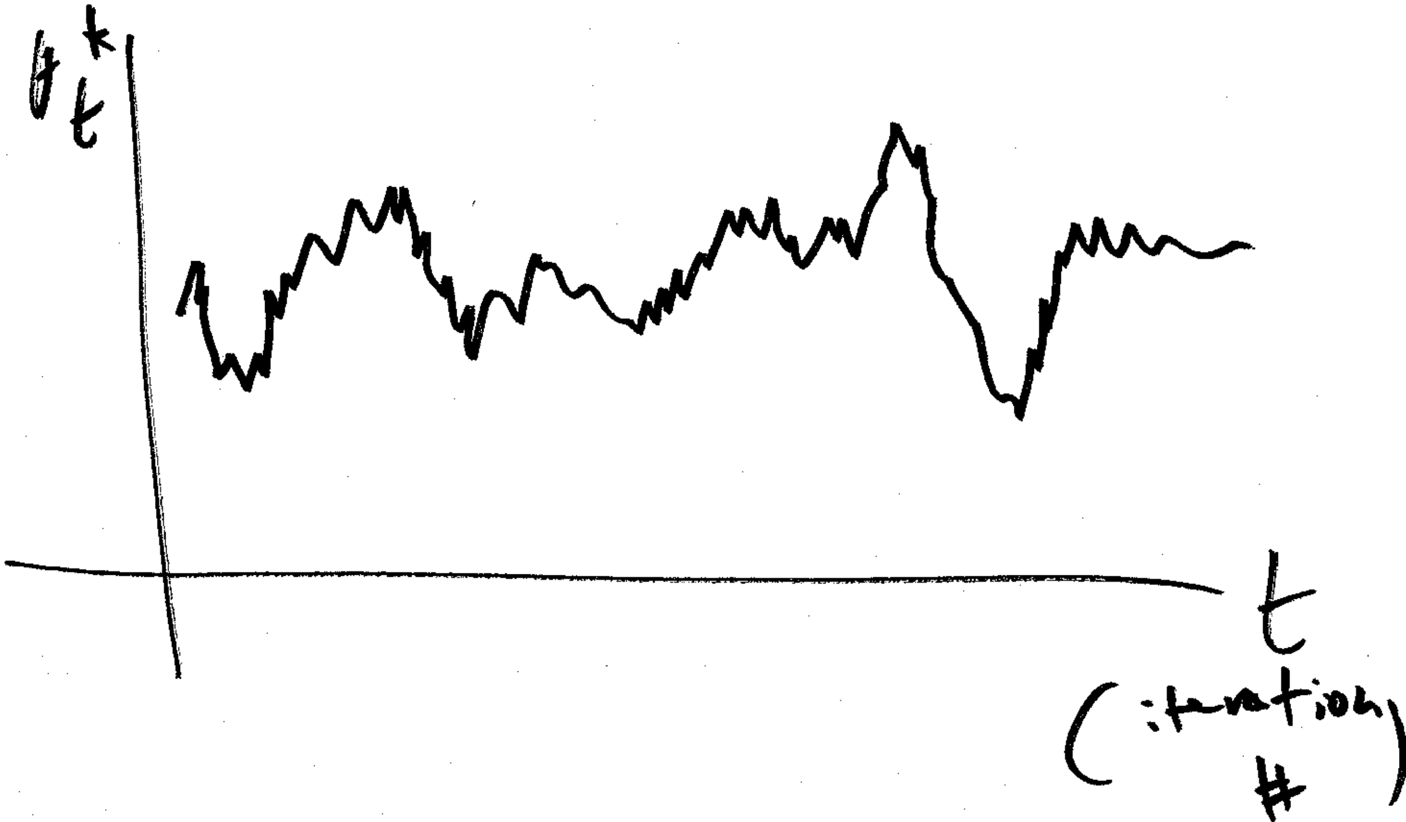



target
p(θ|y)
θ
log p(θ|y)
adaptive rejection sampling (ARS)
log concave

$$
\begin{array}{c|c}
\theta^* & I(\theta^* \le 0.15) \\
\hline
0.17 & 0 \\
0.11 & 1 \\
0.48 & 0 \\
\vdots & \vdots
\end{array}
$$

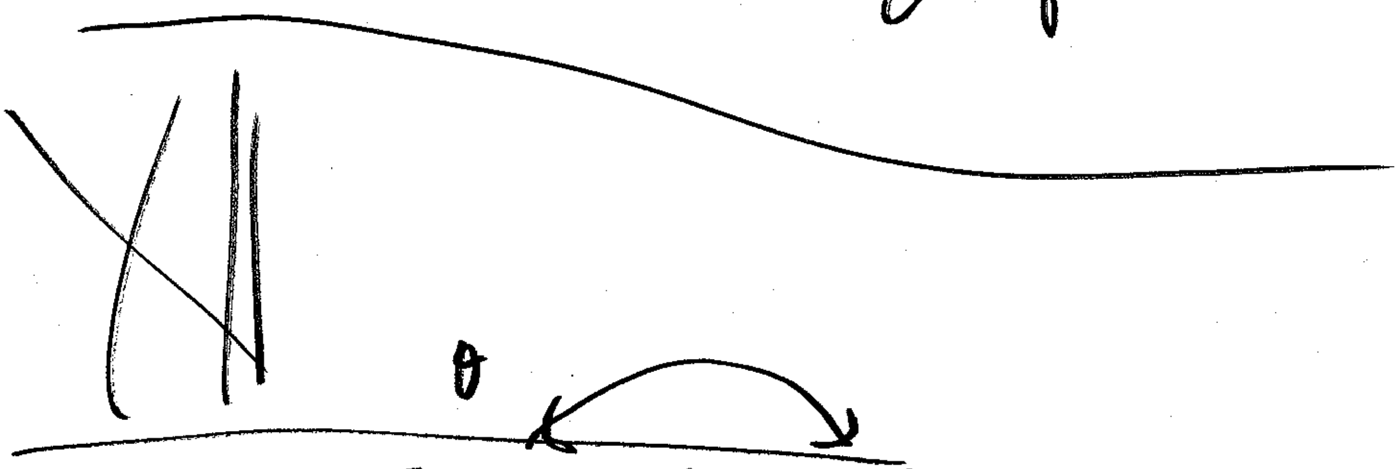
mean $\hat{p}(\theta \le 0.15 \mid y\,B)$

SE $\sqrt{\dfrac{\hat{p}(1-\hat{p})}{m}}$



$$\underbrace{p(\theta \mid y\,\mathcal{B})}_{\text{length } k} = c\,\underbrace{p(\theta \mid \mathcal{B})}\,\underbrace{\ell(\theta \mid y\,\mathcal{B})}$$

c ← hard to compute



monitoring

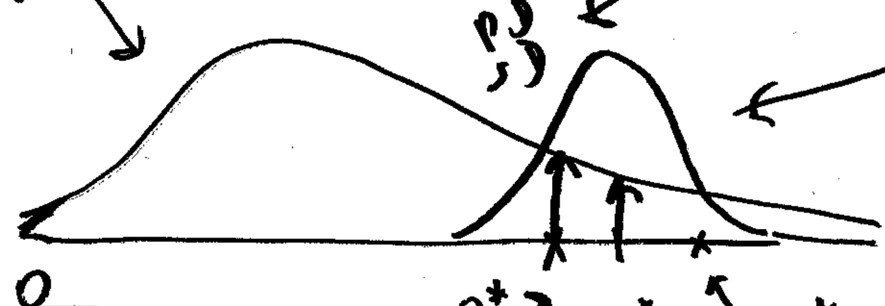
$$\begin{matrix} \theta^*_{(1)} & \theta^*_2 & \theta^*_3 & \cdots \\ \theta^*_2 & \theta^*_3 & \theta^*_4 & \\ \vdots & \vdots & \vdots & \\ \cancel{\theta^*_m} & \cancel{\theta^*_{m+1}} & \theta_{m+2} & \end{matrix} \qquad \hat{\rho}_r$$

$\hat{\rho}_1$ ← MCMC

$\rho_1 \doteq +.8$
$+.9$ $+.95$

target $p(\theta|\cancel{y})$

proposal distribution SD



proposal dist $J(\theta^*|\cancel{y})$

metropolis

I'm thinking of moving here

has to be symmetric

① accept all uphill moves

$$\boxed{\frac{p(\theta^*|y)}{p(\theta^*_t|y)}} > 1$$

② accept any downhill moves with acceptance probability $\dfrac{p(\theta^*|y)}{p(\theta^*_t|y)}$

③ if accept $\theta^*_{t+1} = \theta^*$,
don't accept $\theta^*_{t+1} = \theta^*_t$ (??)

P

$$\left\{ \begin{array}{l} (\mu, \sigma^2, \nu) \sim p(\mu, \sigma^2, \nu) \\ (y_i \mid \mu, \sigma^2, \nu) \stackrel{iid}{\sim} t_\nu(\mu, \sigma^2) \end{array} \right\}$$

NB10

$i = 1, \ldots, n$

$\theta = (\mu, \sigma^2, \nu)$ | $p(\mu, \sigma^2, \nu \mid y\, \mathcal{B})$

complete post.

$p(\mu \mid \sigma^2, \nu, y\, \mathcal{B})$ full cond. for μ

$p(\sigma^2 \mid \mu\, \nu\, y\, \mathcal{B})$ —— σ^2

$p(\nu \mid \mu\, \sigma^2\, y\, \mathcal{B})$ —— ν

$$
\begin{array}{c|c}
 & \begin{array}{ccc} \mu & \sigma^2 & r \end{array} \\
\begin{array}{c} 0 \\ 1 \\ 2 \\ \\ \end{array} & \left[\begin{array}{ccc} \mu_0 & \sigma_0^2 & r_0 \\ \mu_1^* & (\sigma_1^2)^* & r_1^* \\ & & \\ & & \end{array} \right]
\end{array}
$$

draw from $p(\mu \mid \sigma_0^2\ r_0\ y\ \mathcal{B})$

draw from $p(\sigma^2 \mid \mu_1^*\ r_0\ y\ \mathcal{B})$

draw from $p(r \mid \mu_1^*\ (\sigma_1^2)^*\ y\ \mathcal{B})$

always use <u>most</u> <u>recent</u> <u>values</u> of other parameters

---

t dist: RS dist of $\frac{\bar{y} - \mu}{s/\sqrt{n}}$ (F)

$(Y_i \mid \mu, \sigma^2) \overset{iid}{\sim} N(\mu, \sigma^2)$

unknown

t model for data (sampling dist)

normal $\quad Y_i \sim N(\mu, \sigma^2)$

$t_r \quad \left\{ \begin{array}{l} Y_i \overset{indep}{\sim} N(\mu, \sigma_i^2) \\ \sigma_i^2 \overset{iid}{\sim} \text{some dist} \end{array} \right\}$

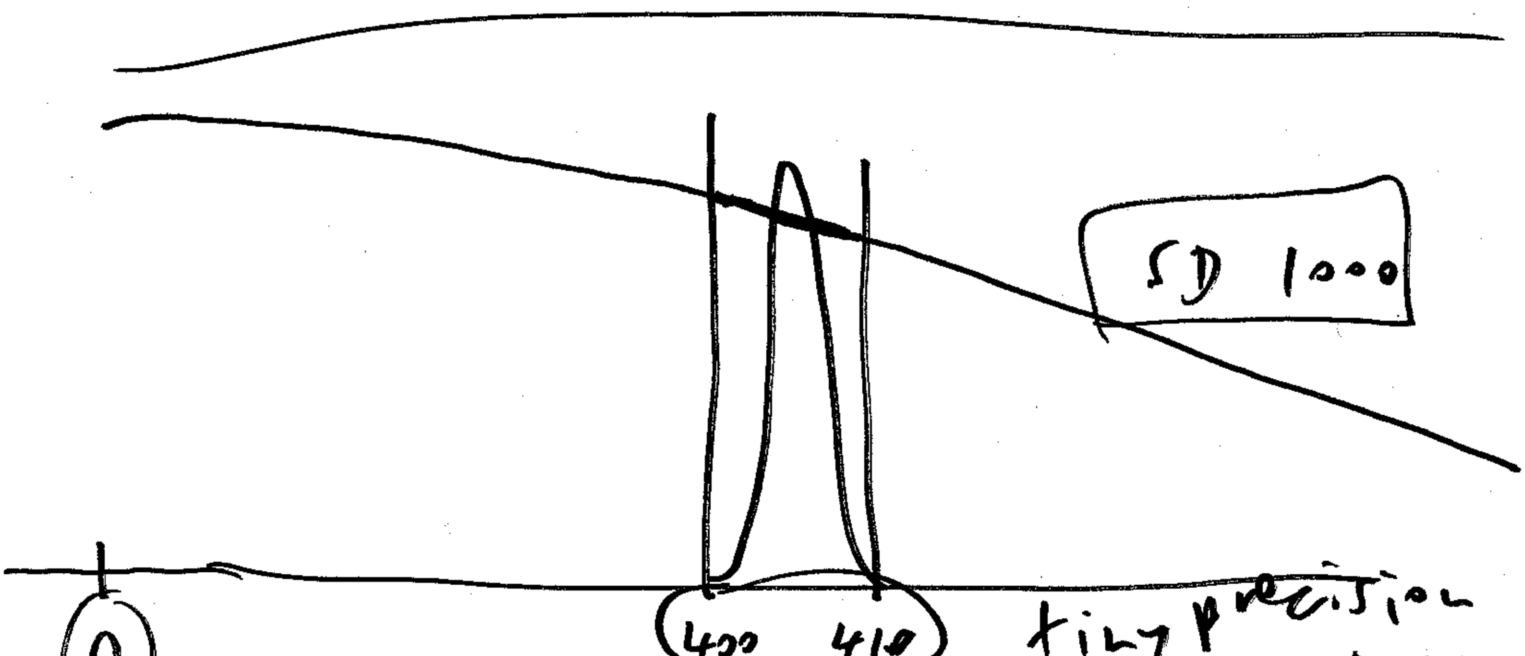
ν

$(\sigma^2 | \cancel{x})$

$(\mu | \sigma^2 \cancel{x})$

$(\lambda_i | \nu \cancel{\sigma^2} \cancel{\mu})$

$(y_i | \lambda_i \cancel{x} \sigma^2 \mu)$

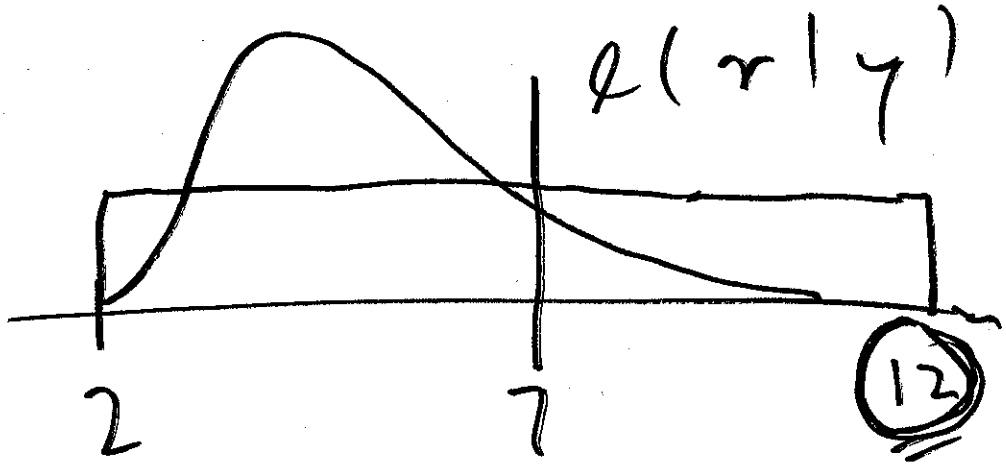


tiny precision

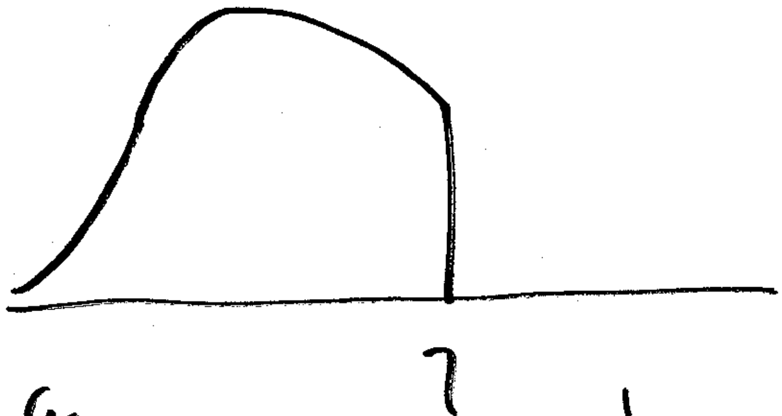
diffuse

$\mu \sim N(0, \text{huge variance})$

1.0 E-6



u(2, e)



MCMC

I.E.D

μ	τ	r	$\sigma = \frac{1}{\sqrt{\tau}}$	Y_{new}
	0.27		$\frac{1}{\sqrt{0.27}}$	✓
	⋮			✓
	⋮			✓
	⋮			

m

$\hat{\text{mean}}$ $E(\mu | y\, B) = \int \mu \, p(\mu | y\, B)\, d\mu$

$p(\mu | y\, B) = \iint p(\mu, \tau, r | y\, B)\, d\tau\, dr$

$$\boxed{p(y_{n+1} \mid \underbrace{y}_{(y_1, \dots, y_n)}\, \mathcal{B})} = \iiint p(y_{n+1}\, \theta \mid y\, \mathcal{B})\, d\theta$$

$$\theta = (\mu, \tau, \nu) \qquad = \iiint p(y_{n+1} \mid \theta\, \cancel{y}\, \mathcal{B}) \cdot p(\theta \mid y\, \mathcal{B})\, d\theta$$

$$= \iiint p(y_{n+1} \mid \theta\, \mathcal{B}) \cdot p(\theta \mid y\, \mathcal{B})\, d\theta$$

$$y \leftrightarrow \left\{ \begin{matrix} x \\ y \mid x \end{matrix} \right\}$$

$$(y_{n+1} \mid y\, \mathcal{B}) \leftrightarrow \left\{ \begin{matrix} (\theta \mid y\, \mathcal{B}) \\ (y_{n+1} \mid \theta\, \cancel{y}\, \mathcal{B}) \end{matrix} \right\}$$

(12)