Math 341 Lec22 5/2/18

1

Model: $X_1, \ldots, X_n \mid \theta_1, \theta_2, \sigma_1^2, \sigma_2^2, \rho \overset{iid}{\sim} \rho N(\theta_1, \sigma_1^2) + (1-\rho) N(\theta_2, \sigma_2^2)$

$P(X_1 \ldots X_n \mid \theta_1, \theta_2, \sigma_1^2, \sigma_2^2, \rho) = \prod_{i=1}^{n}$

the "+" kills us!! Very difficult....

But we can use data augmentation $I_1, \ldots, I_n$.

let $I_1 = 1$ if X_1 belongs to $N(\theta_1, \sigma_1^2)$ and $I_1 = 0$ if X_1 belongs to $N(\theta_2, \sigma_2^2)$

$I_2 = $ "

$\vdots$

$I_n = $ "

$$P(I_1, \ldots, I_n, \theta_1, \theta_2, \sigma_1^2, \sigma_2^2, \rho \mid X_1, \ldots, X_n) \propto P(X_1 \ldots X_n \mid I_1 \ldots I_n, \theta_1, \theta_2, \sigma_1^2, \sigma_2^2, \rho) \cdot P(I_1 \ldots I_n, \theta_1, \theta_2, \sigma_1^2, \sigma_2^2, \rho)$$

$$\underbrace{P(I_1 \ldots I_n \mid \theta_1, \theta_2, \sigma_1^2, \sigma_2^2, \rho)}_{\prod_{i=1}^{n} \rho^{I_i}(1-\rho)^{1-I_i}} \underbrace{P(\theta_1, \theta_2, \sigma_1^2, \sigma_2^2, \rho)}_{\propto 1 \, \propto 1 \, \propto \frac{1}{\sigma_1^2} \, \propto \frac{1}{\sigma_2^2} \, \propto 1}$$

$$= \left(\prod_{i=1}^{n} \left(\frac{1}{\sqrt{2\pi\sigma_1^2}} e^{-\frac{1}{2\sigma_1^2}(X_i - \theta_1)^2} \right)^{I_i} \left(\frac{1}{\sqrt{2\pi\sigma_2^2}} e^{-\frac{1}{2\sigma_2^2}(X_i - \theta_2)^2} \right)^{1-I_i} \rho^{I_i}(1-\rho)^{1-I_i} \right) \frac{1}{\sigma_1^2} \frac{1}{\sigma_2^2}$$

$$= \left(\rho \frac{1}{\sqrt{2\pi\sigma_1^2}} \right)^{\sum I_i} e^{-\frac{1}{2\sigma_1^2} \sum I_i (X_i - \theta_1)^2} \left((1-\rho) \frac{1}{\sqrt{2\pi\sigma_2^2}} \right)^{n - \sum I_i} e^{-\frac{1}{2\sigma_2^2} \sum (1-I_i)(X_i - \theta_2)^2} \frac{1}{\sigma_1^2} \frac{1}{\sigma_2^2}$$

$$P(\theta_1 \mid \ldots\ldots) \propto e^{-\frac{1}{2\sigma_1^2} \sum I_i (X_i^2 - 2X_i\theta_1 + \theta_1^2)} \propto e^{-\frac{1}{2\sigma_1^2}(-2\theta_1 \sum I_i X_i + \theta_1^2 \sum I_i)} = e^{\frac{\sum I_i X_i}{\sigma_1^2}\theta_1 - \frac{\sum I_i}{\sigma_1^2}\theta_1^2}$$

$$\propto N\left(\frac{\sum I_i X_i}{\sum I_i}, \frac{\sigma_1^2}{\sum I_i} \right)$$

$$P(\theta_2 \mid \ldots) \propto N\left(\frac{\sum (1-I_i) X_i}{\sum 1 - I_i}, \frac{\sigma_2^2}{\sum 1 - I_i} \right)$$

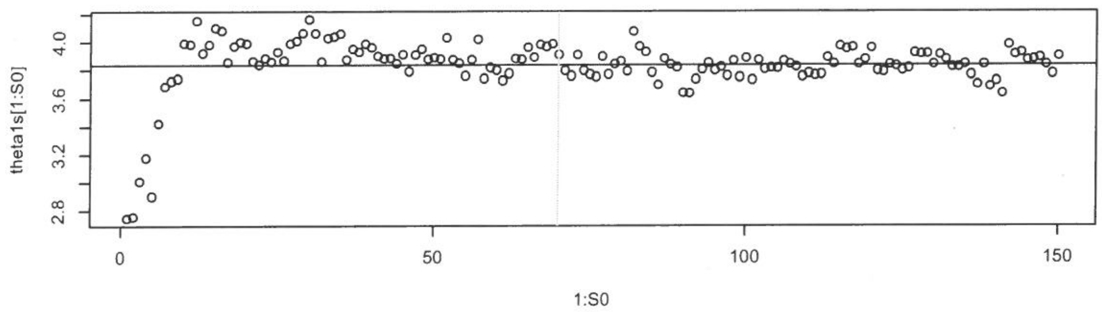
$$P(\sigma_1^2 \mid \ldots) \propto (\sigma_1^2)^{-\frac{\sum I_i}{2} - 1} e^{-\frac{\sum I_i (X_i - \theta_1)^2 / 2}{\sigma_1^2}} \propto \text{InvGamma}\left(\frac{\sum I_i}{2}, \frac{\sum I_i (X_i - \theta_1)^2}{2} \right)$$

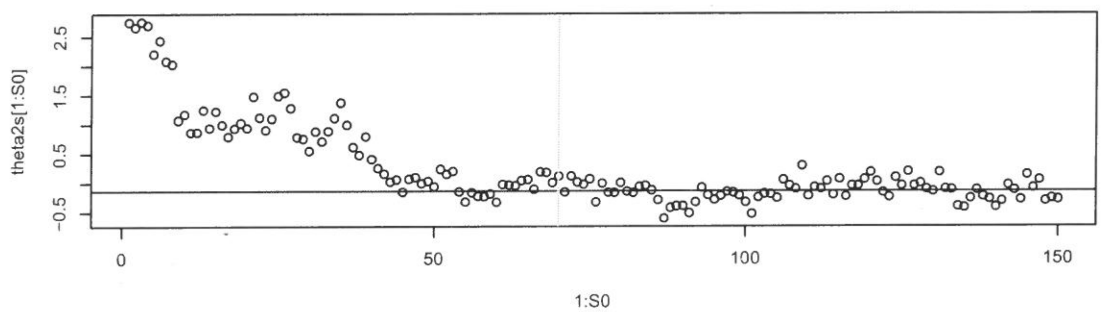
$$P(\sigma_2^2 \mid \ldots) \propto \text{InvG}\left(\frac{\sum 1 - I_i}{2}, \frac{\sum (1 - I_i)(X_i - \theta_2)^2}{2} \right)$$

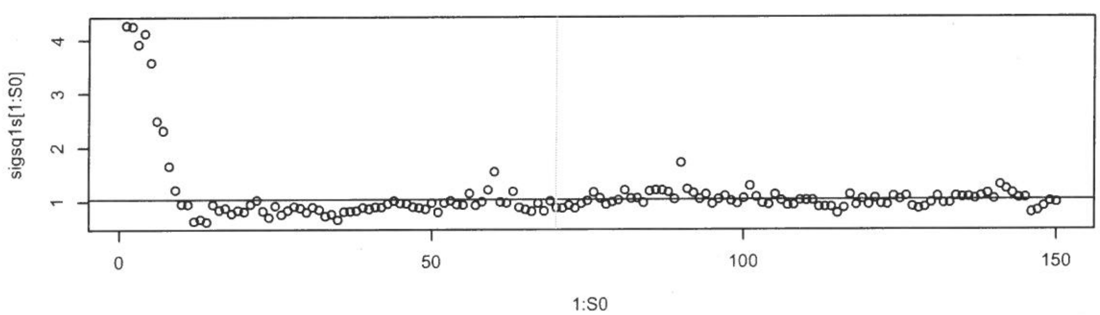
$$P(\rho \mid \ldots) \propto \rho^{\sum I_i} (1-\rho)^{\sum 1 - I_i} \propto \text{Beta}\left(1 + \sum I_i, 1 + \sum 1 - I_i \right)$$

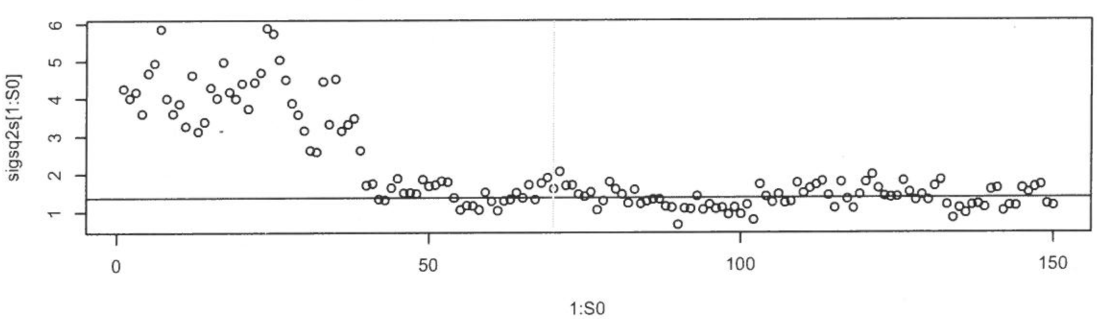
$$P(I_i \mid \ldots) \propto \Big(\underbrace{\rho \frac{1}{\sqrt{2\pi\sigma_1^2}} e^{-\frac{1}{2\sigma_1^2}(X_i - \theta_1)^2}}_{a} \Big)^{I_i} \Big(\underbrace{(1-\rho) \frac{1}{\sqrt{2\pi\sigma_2^2}} e^{-\frac{1}{2\sigma_2^2}(X_i - \theta_2)^2}}_{b} \Big)^{1-I_i} \propto \text{Bern}\left(\frac{a}{a+b} \right)$$

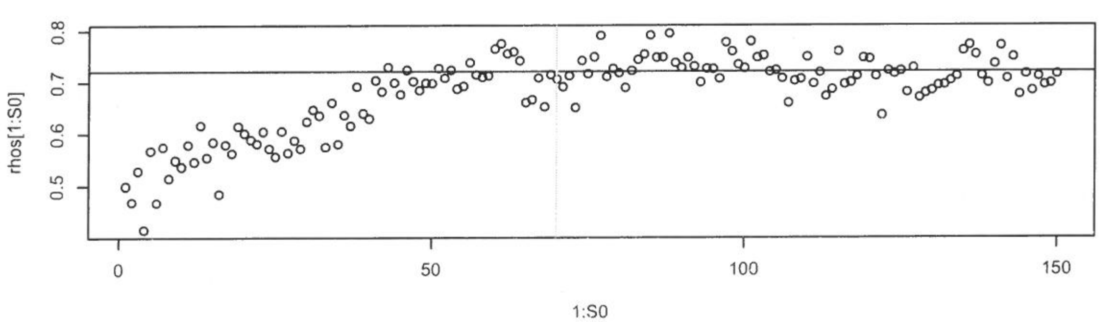
$\forall i$


theta1s[1:S0]
2.8
3.2
3.6
4.0
0
50
100
150
1:S0


theta2s[1:S0]
-0.5
0.5
1.5
2.5
0
50
100
150
1:S0

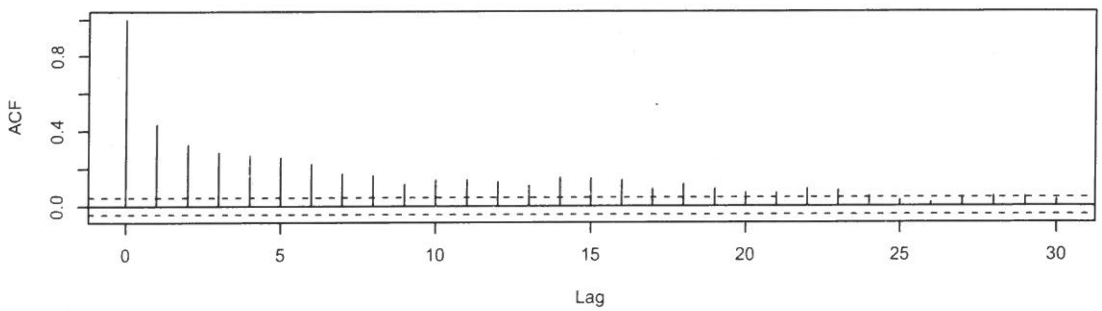

sigsq1s[1:S0]
1
2
3
4
0
50
100
150
1:S0


sigsq2s[1:S0]
1
2
3
4
5
6
0
50
100
150
1:S0

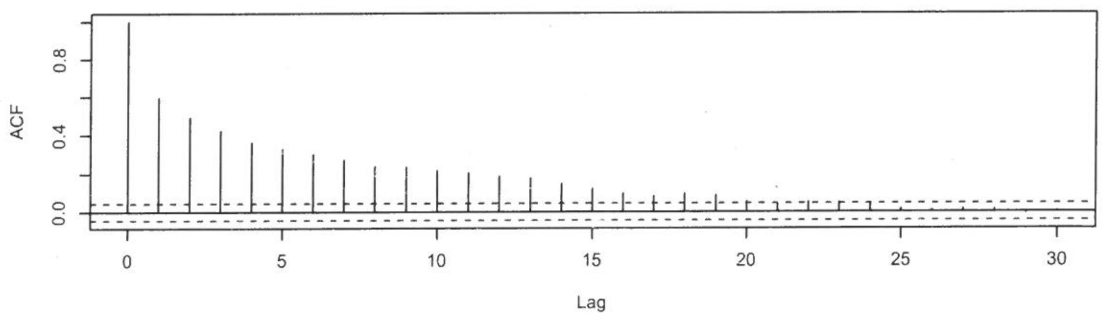

rhos[1:S0]
0.5
0.6
0.7
0.8
0
50
100
150
1:S0

It's
[illegible]!

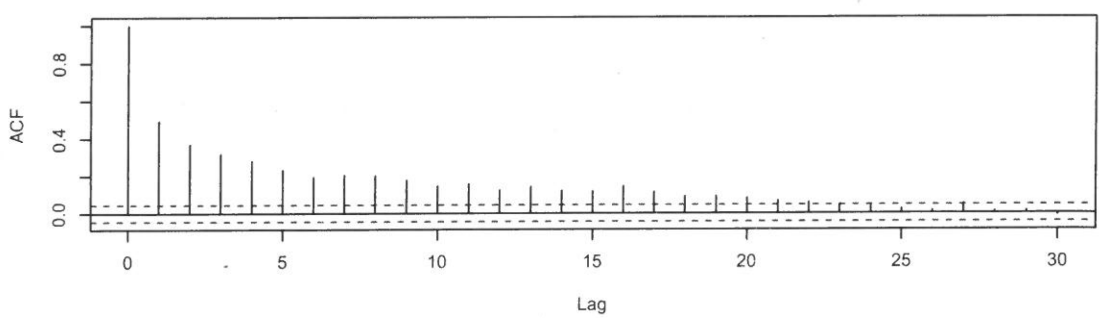
Series theta1s[B:S]



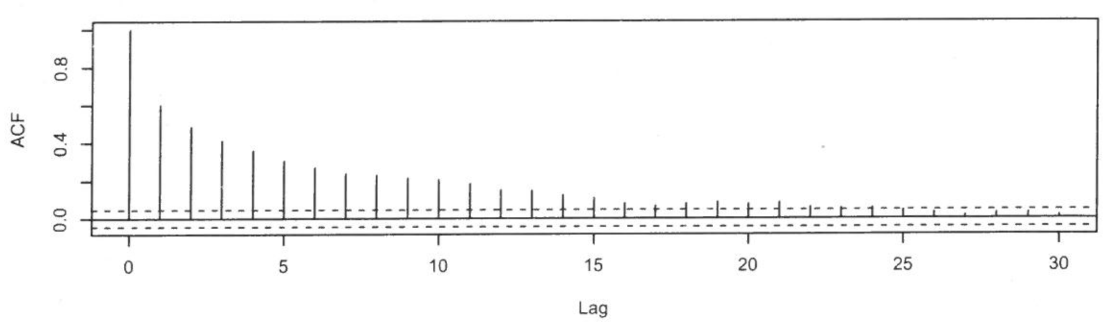
Series theta2s[B:S]



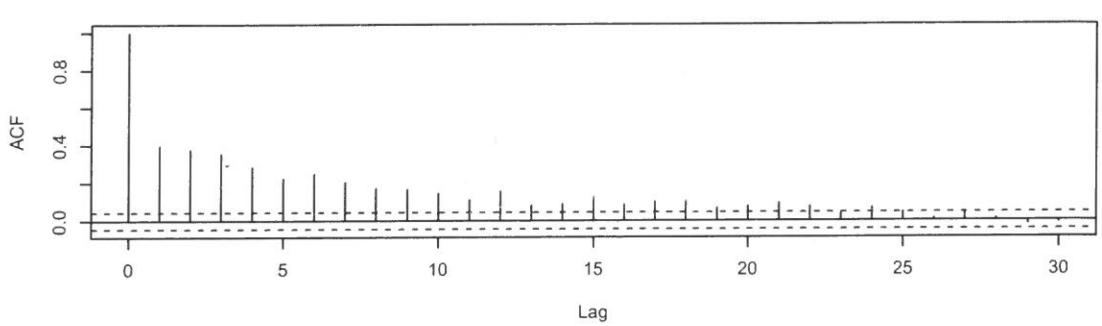
Series sigsq1s[B:S]



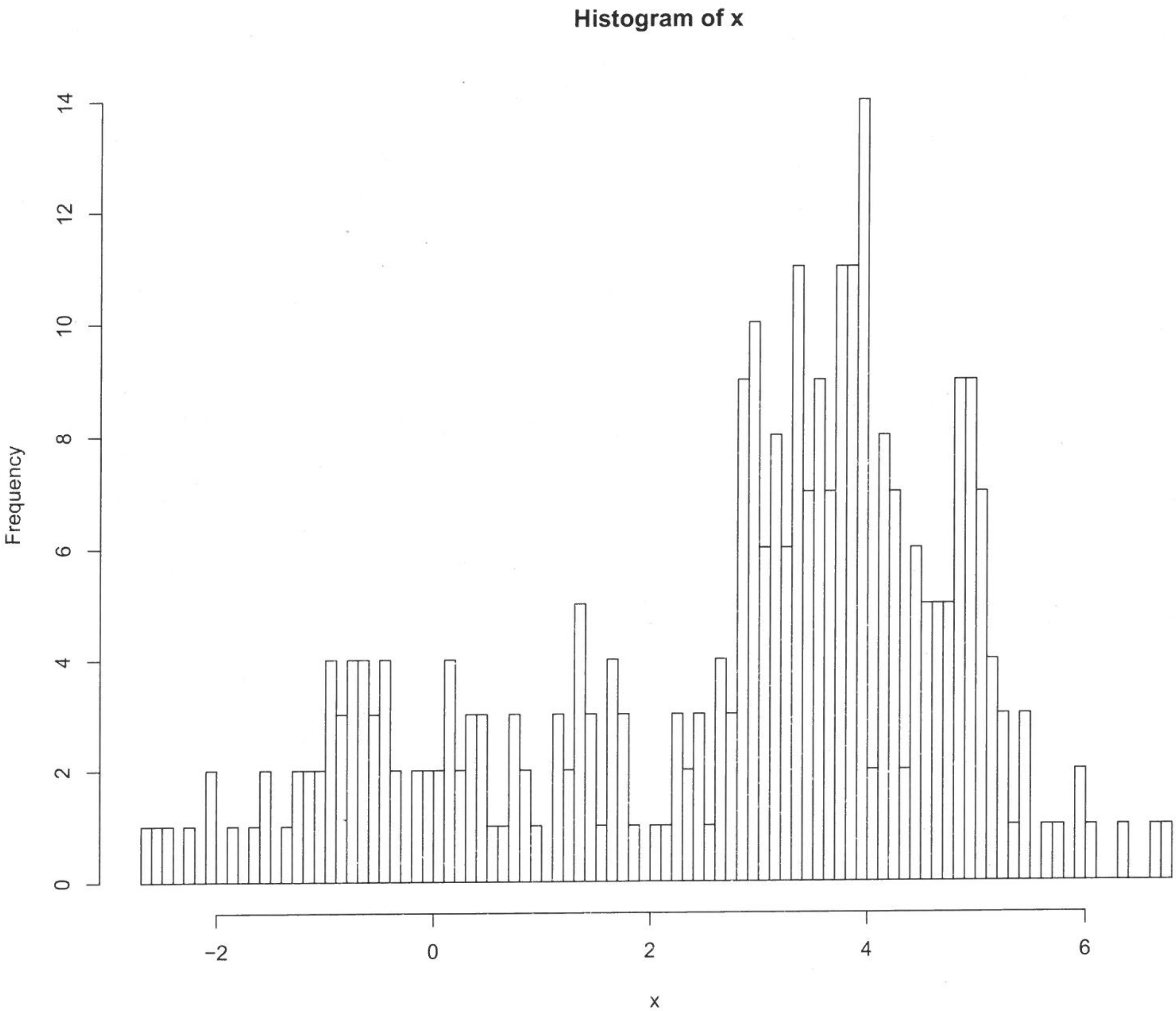
Series sigsq2s[B:S]

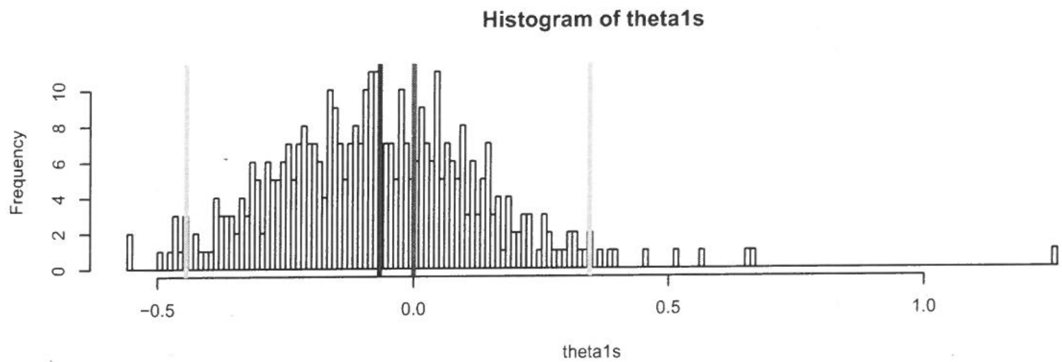


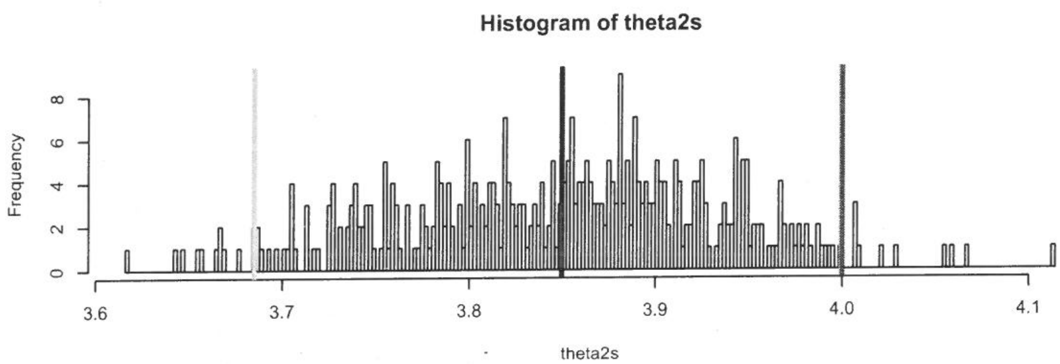
Series rhos[B:S]

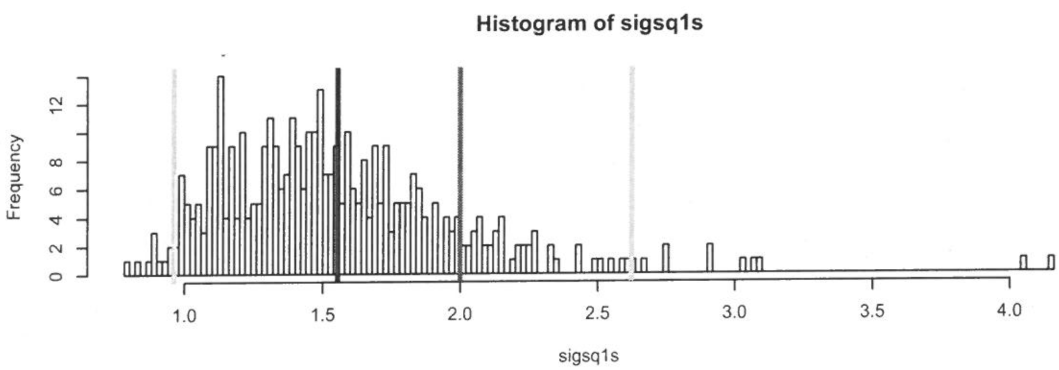


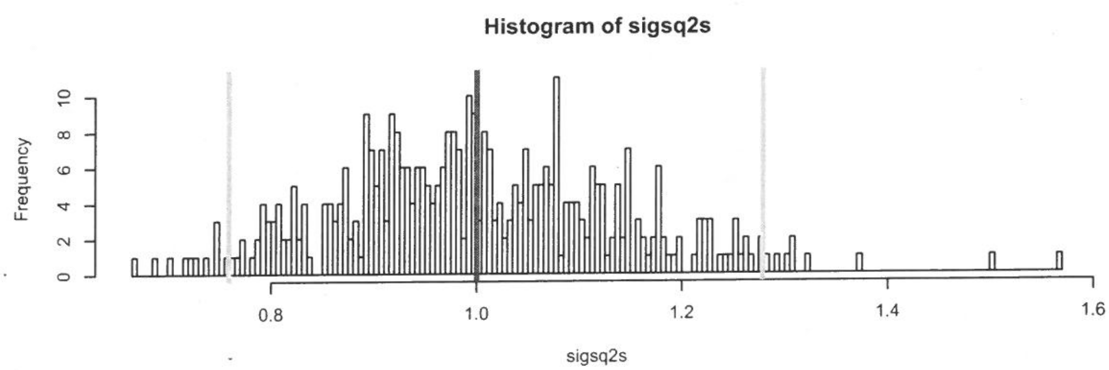
Mixture Model Gibbs Sampler

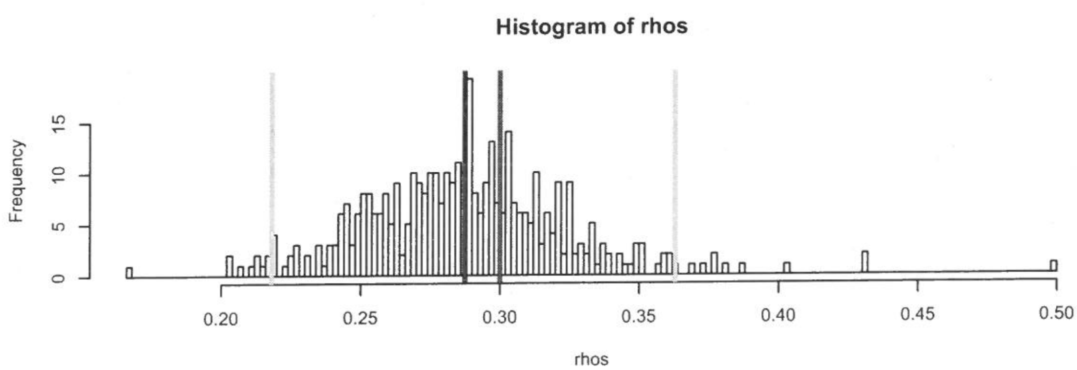



Histogram of theta1s
Frequency
0
2
4
6
8
10
−0.5
0.0
0.5
1.0
theta1s


Histogram of theta2s
Frequency
0
2
4
6
8
3.6
3.7
3.8
3.9
4.0
4.1
theta2s


Histogram of sigsq1s
Frequency
0
2
4
6
8
12
1.0
1.5
2.0
2.5
3.0
3.5
4.0
sigsq1s


Histogram of sigsq2s
Frequency
0
2
4
6
8
10
0.8
1.0
1.2
1.4
1.6
sigsq2s


Histogram of rhos
Frequency
0
5
10
15
0.20
0.25
0.30
0.35
0.40
0.45
0.50
rhos

$$P(Q) \propto \underbrace{0.1^{Q}\, 0.2^{1-Q}}_{= k(Q)} \qquad \text{Supp}(Q) = \{0,1\}$$

$= ?$

$$P(Q=1) = \frac{k(1)}{k(0)+k(1)} = \frac{0.1}{0.1+0.2}$$

$P(I_2 \mid \cdots) = \text{same}$

$\vdots$

$P(I_n \mid \cdots) = \text{same}$

Metropolis-Hastings Algorithm

Recall change pt model...

$\overset{iid}{\sim} \text{Poisson}(\lambda_1)$ $\qquad$ $\overset{iid}{\sim} \text{Poisson}(\lambda_2)$

What if instead...

$$P(X_t) = \text{Poisson}(\overset{\lambda}{\overbrace{a + bt}})$$

i.e. λ is a function of time, and dependent

a linear function of time.

this is called a "Poisson Regression"

e.g. the dance sec...

t	X
1	3
1.3	0
2	5
2.7	7
3.1	6
4.9	5
5.0	5
7.5	9

the likelihood

$$P(X|a,b) = \prod_{i=1}^{n} \frac{(a+bt_i)^{x_i} e^{-(a+bt_i)}}{x_i!} = \frac{\left(\prod (a+bt_i)^{x_i}\right) e^{-\sum a+bt_i}}{\prod x_i!}$$

You care about λ i.e. you care about finding a, b; $\propto e^{-\sum a+bt_i} \prod (a+bt_i)^{x_i}$

$$P(a,b|X) \propto P(X|a,b)\, P(a,b) = P(X|a,b)$$

prior? $P(a) \propto 1$, $P(b) \propto 1$

$a, b \in \mathbb{R}$

$$\Rightarrow P(a,b|x) \propto e^{-\sum a+bt_i} \prod (a+bt_i)^{x_i}$$ non-conj.

Gibbs sampling to the rescue?

$P(a|X,b) \propto$ ditto

$P(b|X,a) \propto$ ditto

In order to use Gibbs sampling, you would have to grid sample $k(a|X,b)$ and $k(b|X,a)$ ... TOO SLOW!!!

Need something else ... Metropolis-Hastings Algorithm

E. Teller!

Metropolis et al 1953, Hastings, 1970

What does Gibbs sampling do? It moves around the space...

Why not move around another way?