Lec 21 Math 341 5/7/18

If we prove

$$P(x) = \lim_{t \to \infty} \int_X \left(\prod_{i=0}^{t+1} P(x_i \mid x_{i-1}) \right) P(X_0 = x)\, dx$$

the MC converges

Thm: if $P(X_1, \ldots, X_p)$ hold the positivity condition then...

$$\forall \vec{a} \in \mathcal{X}, \quad P(X_1, \ldots X_p) \propto \prod_{j=1}^{p} \frac{P(X_j \mid X_1, \ldots X_{j-1}, X_{j+1} = a_{j+1}, \ldots, X_p = a_p)}{P(X_j = a_j \mid X_1, \ldots, X_{j-1}, X_{j+1} = a_{j+1}, \ldots, X_p = a_p)}$$

Back to Gibbs sampler... the transition kernel is:

$$P(\vec{\theta}_{t+1} \mid \vec{\theta}_t, X) = P(\theta_{t+1,1}, \ldots, \theta_{t+1,p} \mid \theta_{t,1}, \ldots \theta_{t,p}, X)$$

I will drop the X notation for convenience going forward

Systematic sweep steps ... NOT Bayes Rule !!

$$= P(\theta_{t+1,p} \mid \theta_{t+1,1}, \dots, \theta_{t+1,p-1}) \cdot$$

not t+1!

$$P(\theta_{t+1,p-1} \mid \theta_{t+1,1}, \dots, \theta_{t+1,p-2}, \theta_{t,p}) \cdot$$

$$P(\theta_{t+1,p-2} \mid \theta_{t+1,1}, \dots, \theta_{t+1,p-3}, \theta_{t,p-1}, \theta_{t,p-2}) \cdot$$

$$\vdots \quad \cdot$$

$$P(\theta_{t+1,2} \mid \theta_{t+1,1}, \theta_{t,3}, \dots, \theta_{t,p}) \cdot$$

$$P(\theta_{t+1,1} \mid \theta_{t,2}, \dots, \theta_{t,p})$$

represents the P steps

$$P(\vec{\theta}_{t+1}) = \int_{\Theta} P(\vec{\theta}_{t+1} \mid \vec{\theta}_t)\, P(\vec{\theta}_t)\, d\vec{\theta}$$

if $P(\vec{\theta}_{t+1}) = P(\vec{\theta}_t)$ then this distr is the invariant distr

$$P(\theta_{t+1,1}, \theta_{t+1,2}, \dots, \theta_{t+1,p}) = \int_{\theta_{t,1}}\int_{\theta_{t,2}} \cdots \int_{\theta_{t,p}} P(\theta_{t,1}, \theta_{t,2}, \dots, \theta_{t,p})\, d\theta_{t,1}\, d\theta_{t,2} \cdots d\theta_{t,p}$$

$$\| $$

$$P(\theta_{t,1} \mid \theta_{t,2}, \dots, \theta_{t,p})\, P(\theta_{t,2}, \dots, \theta_{t,p})$$

$$= \int_{\theta_{t,2}} \cdots \int_{\theta_{t,p}} P(\theta_{t,2}, \dots \theta_{t,p}) \underbrace{\int_{\theta_{t,1}} P(\theta_{t,1} \mid \theta_{t,2}, \dots \theta_{t,p})\, d\theta_{t,1}}_{1} \cdot d\theta_{t,2} \cdots d\theta_{t,p}$$

$$= \int_{\theta_{t,3}} \cdots \int_{\theta_{t,p}}^{\text{rest of kernel}} \underbrace{\int_{\theta_{t,2}} \underbrace{P(\theta_{t+1,1} \mid \theta_{t,2}, \dots, \theta_{t,p})\, P(\theta_{t,2}, \dots \theta_{t,p})}_{P(\theta_{t+1,1}, \theta_{t,2}, \dots, \theta_{t,p})} d\theta_{t,2}}_{P(\theta_{t+1,1}, \theta_{t,3}, \dots \theta_{t,p})}\, d\theta_{t,3} \cdots d\theta_{t,p}$$

$$\int_{\theta_{t,4}} \cdots \int_{\theta_{t,p}} \overset{\text{rest of terms}}{\cdots\cdots} \underbrace{\int_{\theta_{t,3}} P(\theta_{t+1,2} \mid \theta_{t+1,1}, \theta_{t,3}, \dots, \theta_{t,p})\, P(\theta_{t+1,1}, \theta_{t,3}, \dots, \theta_{t,p})\, d\theta_{t,3}}_{P(\theta_{t+1,2},\, \theta_{t+1,1},\, \theta_{t,4}, \dots, \theta_{t,p})}\, d\theta_{t,4} \cdots d\theta_{t,p}$$

$$\vdots$$

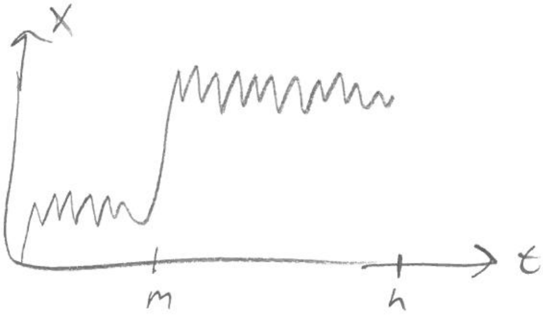
$$P(\theta_{t+1,1}, \dots \theta_{t+p,p})$$

proves that the Gibbs Sampler converges.

Some examples...

Suppose you want to fit some advanced models...

Change pt model



$\overset{iid}{\sim}$ Poisson(λ_1) $\quad\overset{iid}{\sim}$ Poisson(λ_2)

Both λ_1, λ_2 unknown and m (the change pt) unknown.

Priors

$\lambda_1 \sim \text{Gamma}(\alpha, \beta)$, $\lambda_2 \sim \text{Gamma}(\alpha, \beta)$, $P(m) \propto \frac{1}{n}$ i.e. Uniform discrete

$$P(\lambda_1, \lambda_2, m \mid X_1, \ldots, X_n) \propto P(X_1, \ldots, X_n \mid \lambda_1, \lambda_2, m)\, P(\lambda_1)\, P(\lambda_2)\, P(m)$$

$$\propto \left(\prod_{i=1}^{m} \frac{e^{-\lambda_1} \lambda_1^{x_i}}{x_i!}\right)\left(\prod_{i=m+1}^{n} \frac{e^{-\lambda_2} \lambda_2^{x_i}}{x_i!}\right) \lambda_1^{\alpha-1} e^{-\beta\lambda_1} \lambda_2^{\alpha-1} e^{-\beta\lambda_2}$$

$$\propto e^{-m\lambda_1} \lambda_1^{\sum_{i=1}^{m} x_i} e^{-(n-m)\lambda_2} \lambda_2^{\sum_{i=m+1}^{n} x_i} \lambda_1^{\alpha-1} e^{-\beta\lambda_1} \lambda_2^{\alpha-1} e^{-\beta\lambda_2}$$

$$= e^{-(m+\beta)\lambda_1} \lambda_1^{\sum_{i=1}^{m} x_i + \alpha - 1} e^{-(n-m+\beta)\lambda_2} \lambda_2^{\sum x_i + \alpha - 1}$$

non-conj!

$$P(\lambda_1 \mid X_1, \ldots, X_n, \lambda_2, m) \propto \text{Gamma}\left(\sum_{i=1}^{m} x_i + \alpha,\ m + \beta\right)$$

$$P(\lambda_2 \mid X_1, \ldots, X_n, \lambda_1, m) \propto \text{Gamma}\left(\sum_{i=m+1}^{n} x_i + \alpha,\ n - m + \beta\right)$$

$$P(m \mid X_1, \ldots, X_n, \lambda_1, \lambda_2) \propto e^{-m(\lambda_1 - \lambda_2)} \lambda_1^{\sum x_i} \lambda_2^{\sum x_i} = k(m \mid \cdot)$$

$$= \frac{k(m \mid \cdot)}{\sum_{m=1}^{n} k(m \mid \cdot)}$$

since uniform discrete with few elements in its support, it's very easy to sample from using grid sampling

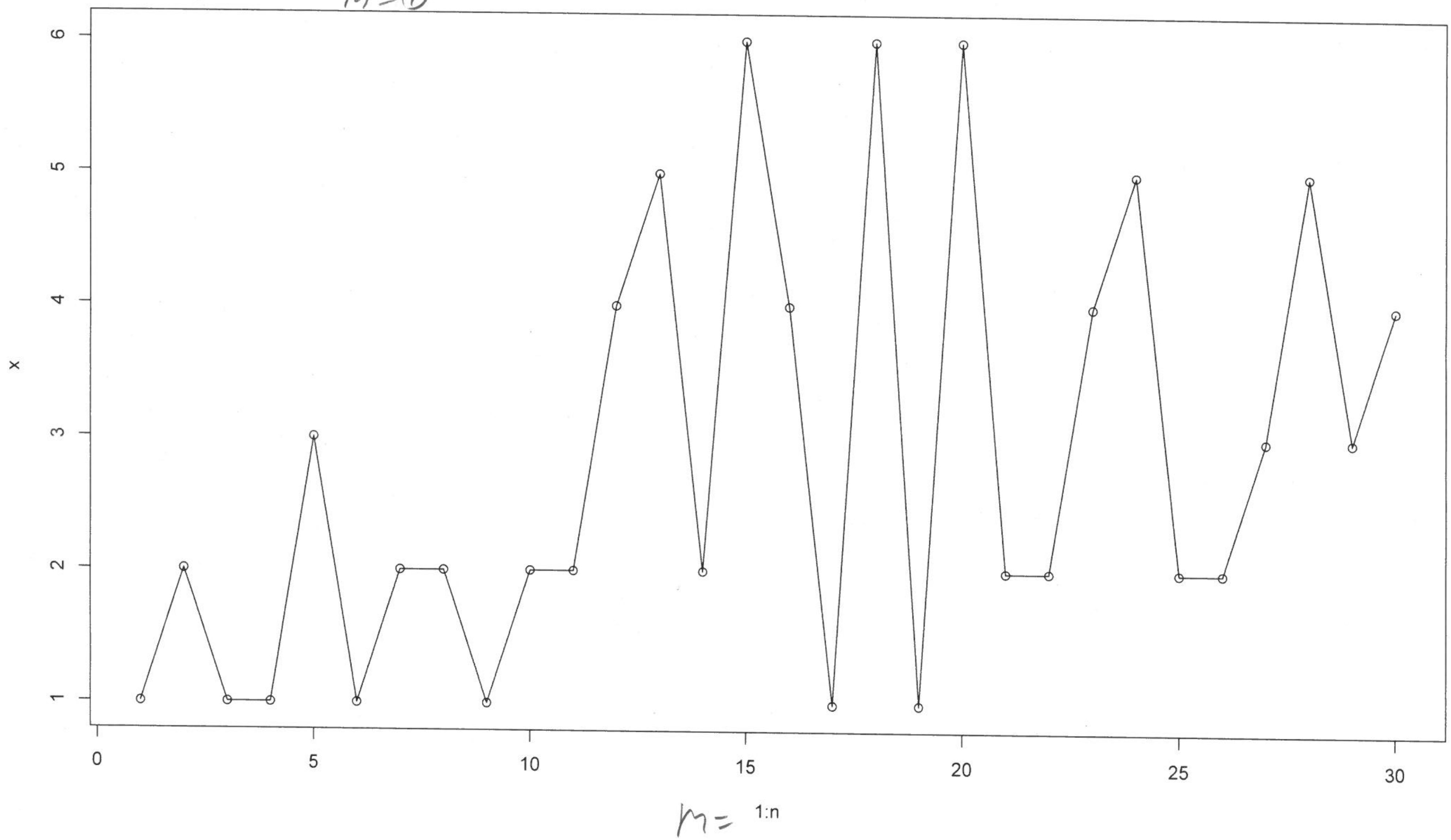
Change pt Poisson Model

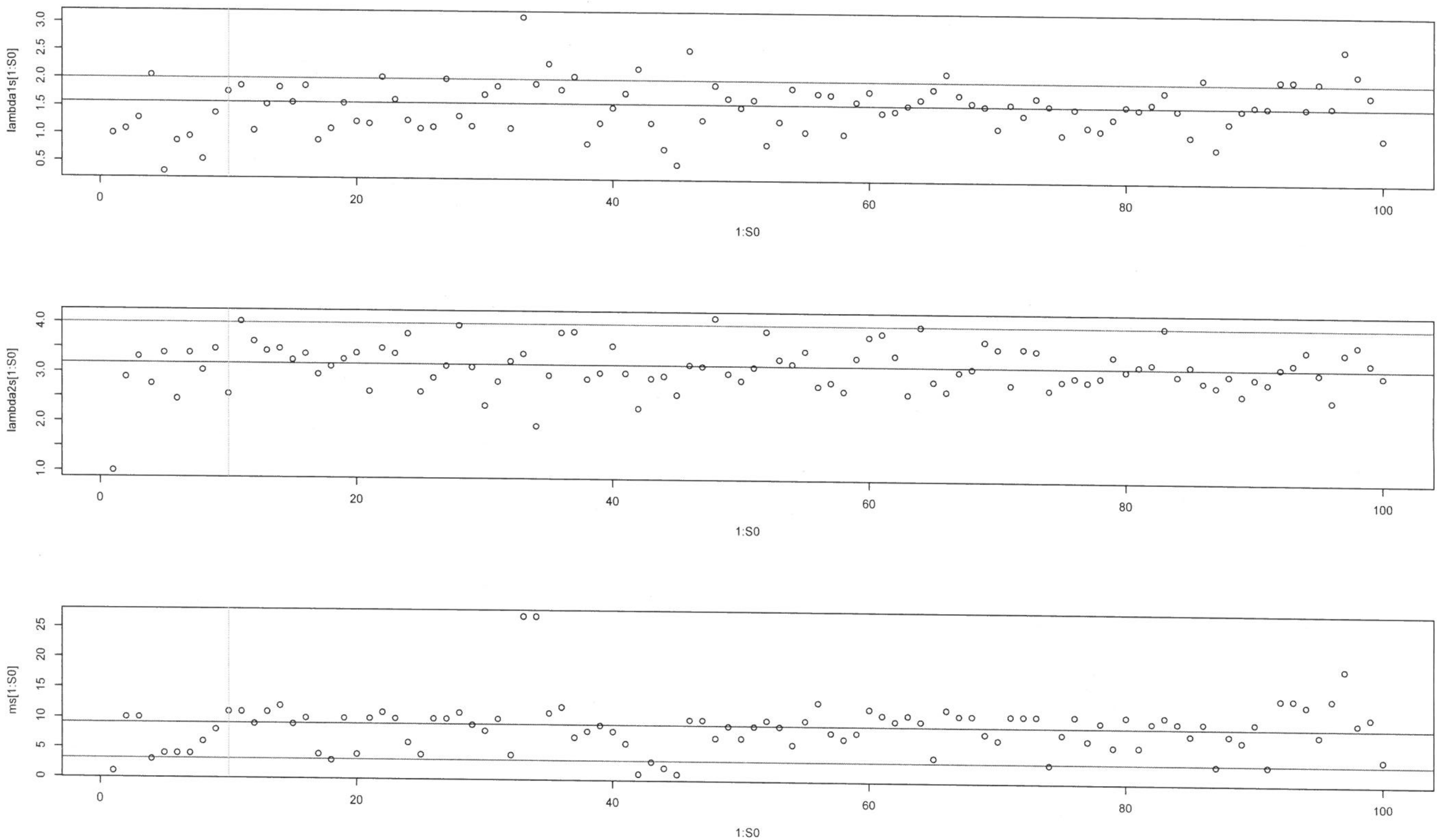
true param value

$\lambda_1 = 2$

$\lambda_2 = 4$

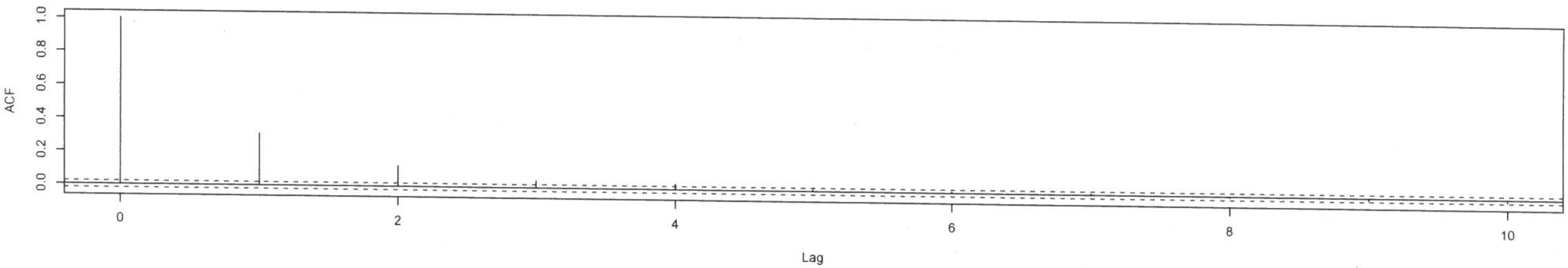
$m = 10$



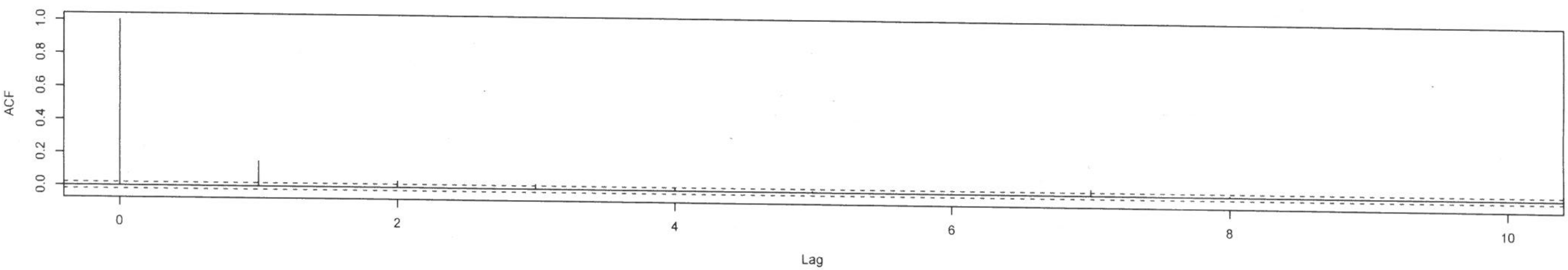


lambda1s[1:S0]
3.0 2.5 2.0 1.5 1.0 0.5
0 20 40 60 80 100
1:S0
lambda2s[1:S0]
4.0 3.0 2.0 1.0
0 20 40 60 80 100
1:S0
ms[1:S0]
25 20 15 10 5 0
0 20 40 60 80 100
1:S0

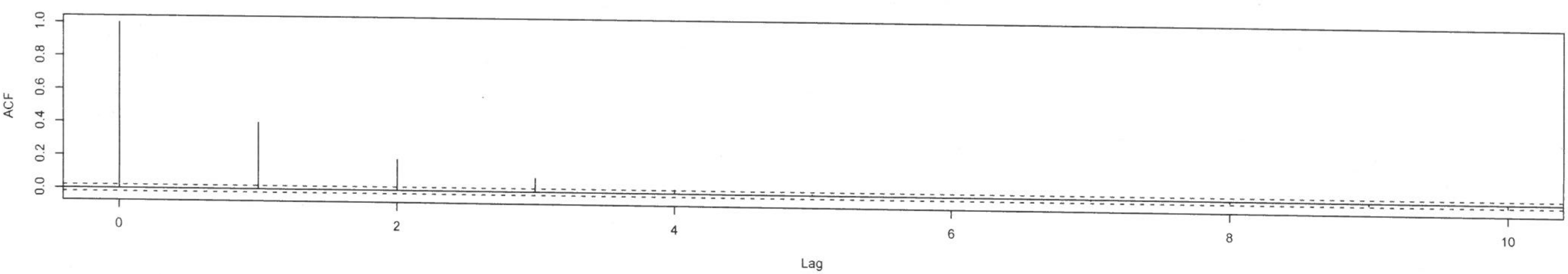
Series lambda1s[B:S]



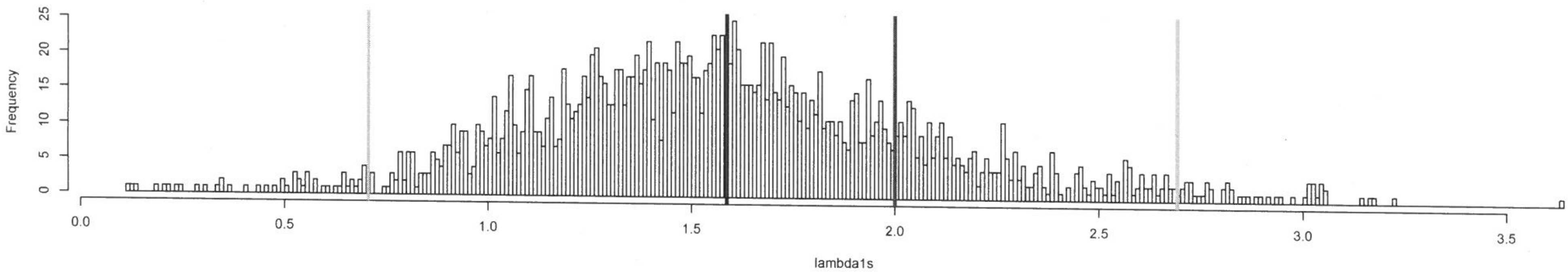
Series lambda2s[B:S]



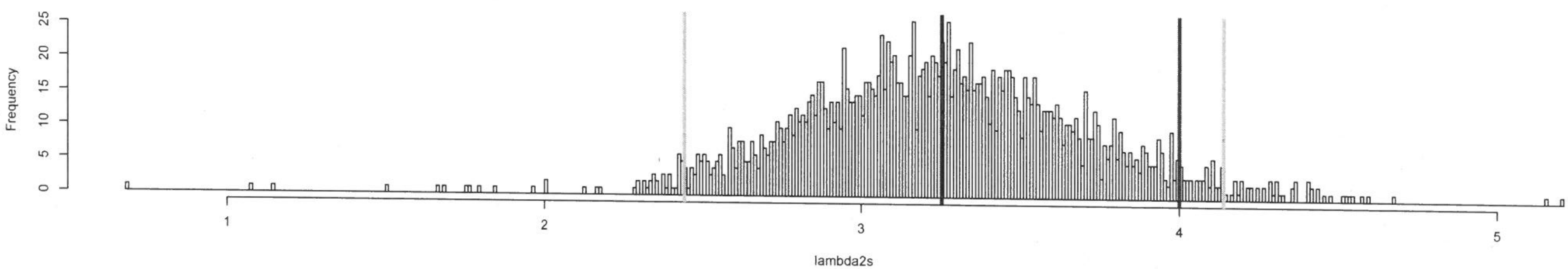
Series ms[B:S]



Histogram of lambda1s



Histogram of lambda2s



Histogram of ms