Lecture 8

- Markov Chain
- Description on how a MCMC algorithm works
- diagnostic of a MCMC
- exercises

Markov Chain

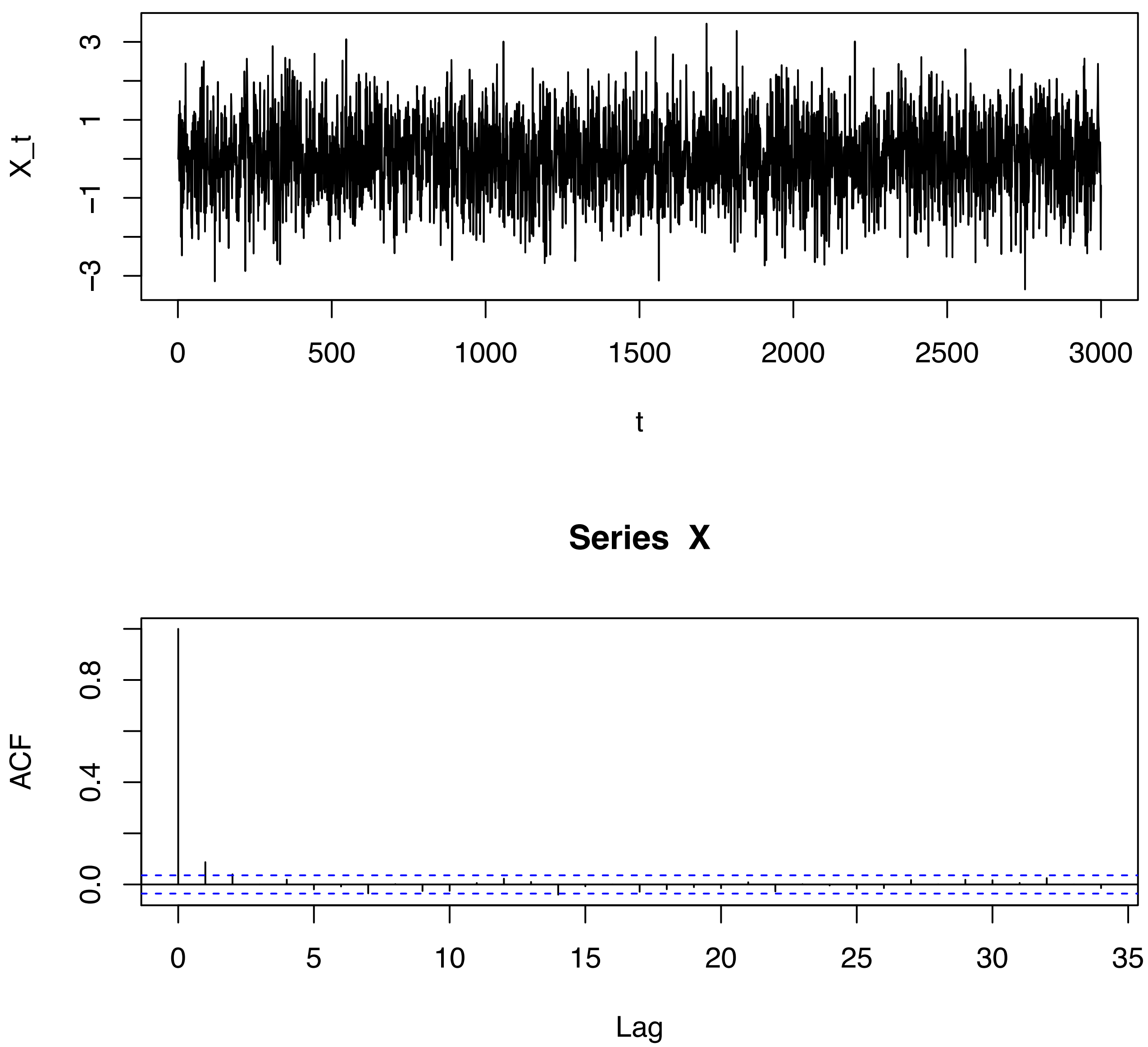
Markov property, (or memory less property):

$$P(X_t < c | X_1, X_2, \ldots, X_{t-1}) = P(X_t < c | X_{t-1})$$

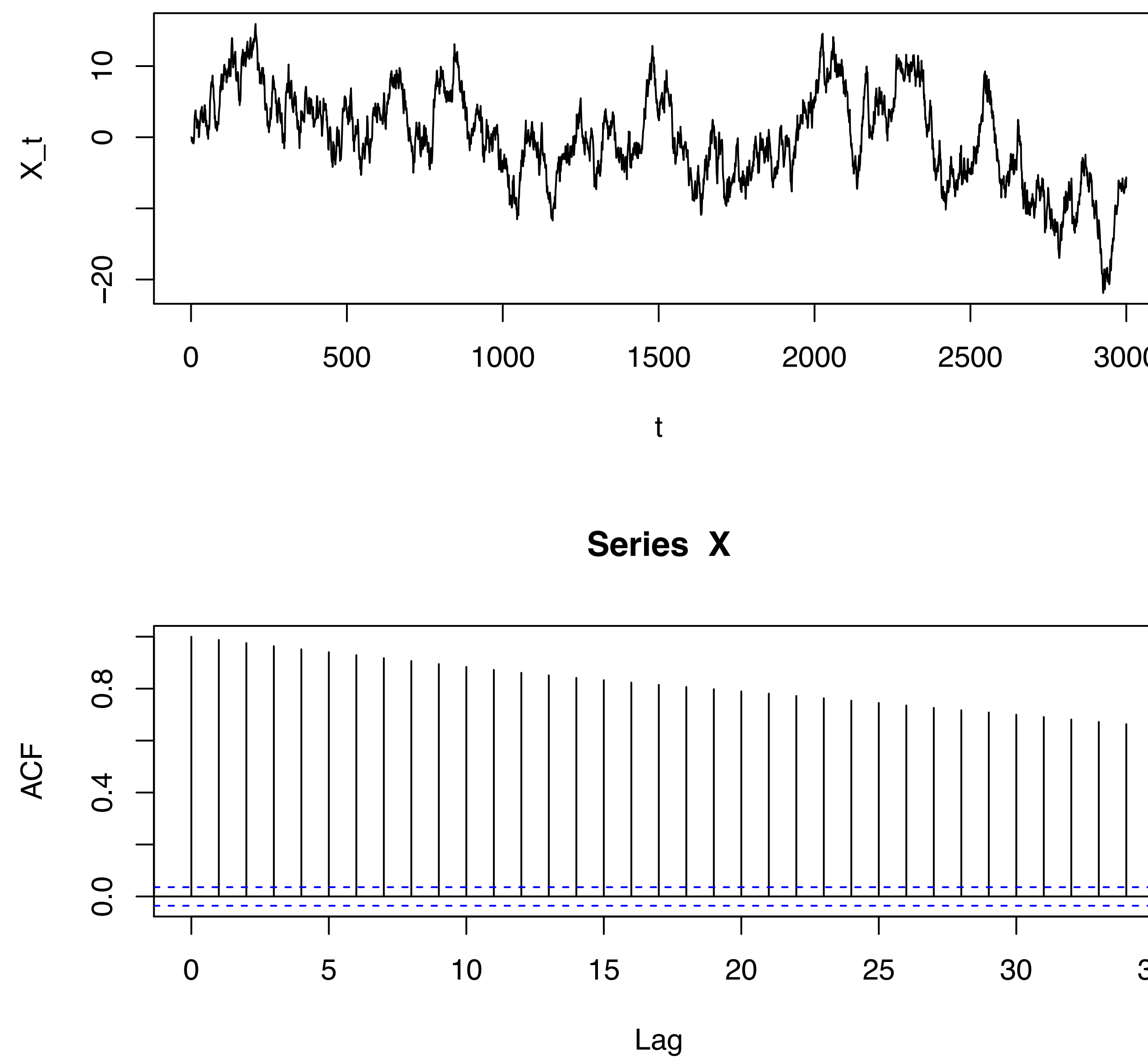
example: AR(1)

$$X_t = aX_{t-1} + \epsilon_t$$

a=0.1


X_t
3
1
−1
−3
0
500
1000
1500
2000
2500
3000
t
Series X
ACF
0.8
0.4
0.0
0
5
10
15
20
25
30
35
Lag

a=0.9

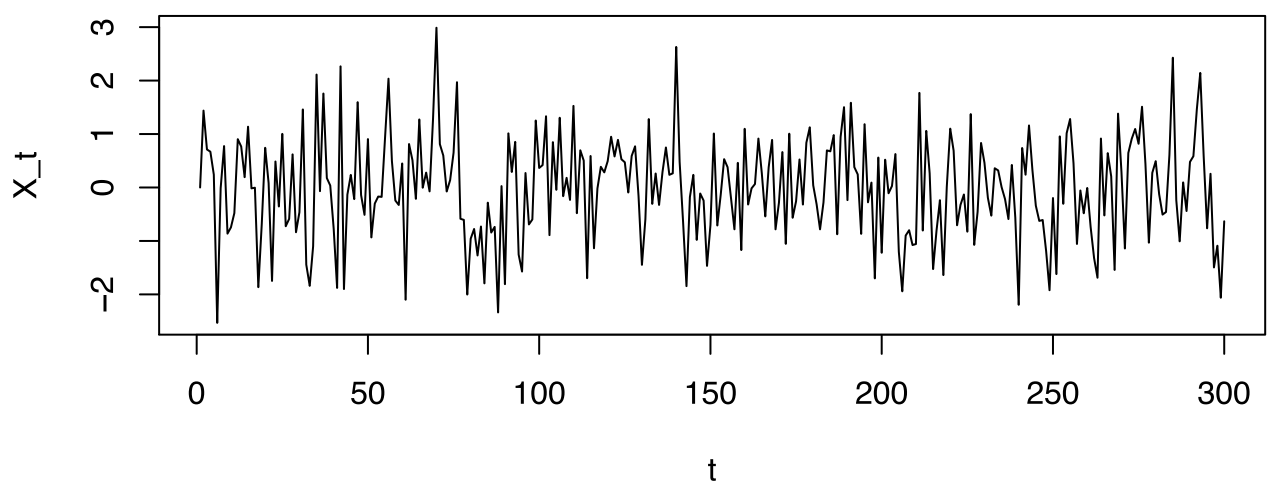

X_t
10
0
−20
0
500
1000
1500
2000
2500
3000
t
Series X
ACF
0.8
0.4
0.0
0
5
10
15
20
25
30
35
Lag

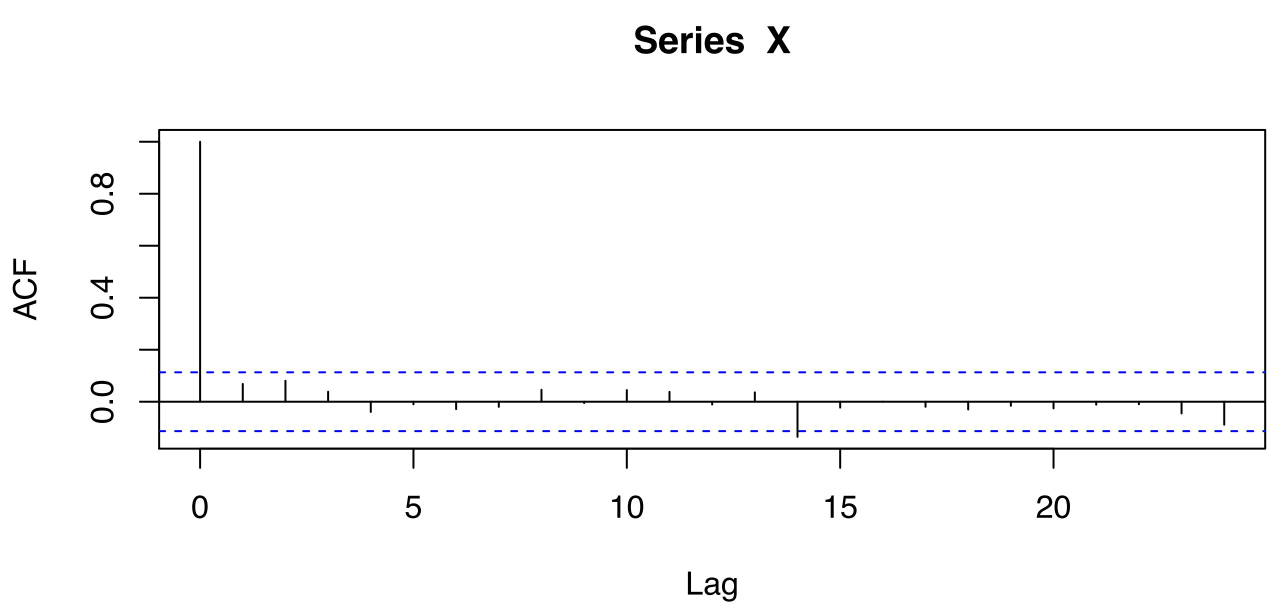
For well behaved Markov chains
if one takes a large T then the sequence

$$X_T, X_{2T}, X_{3T}, \ldots$$

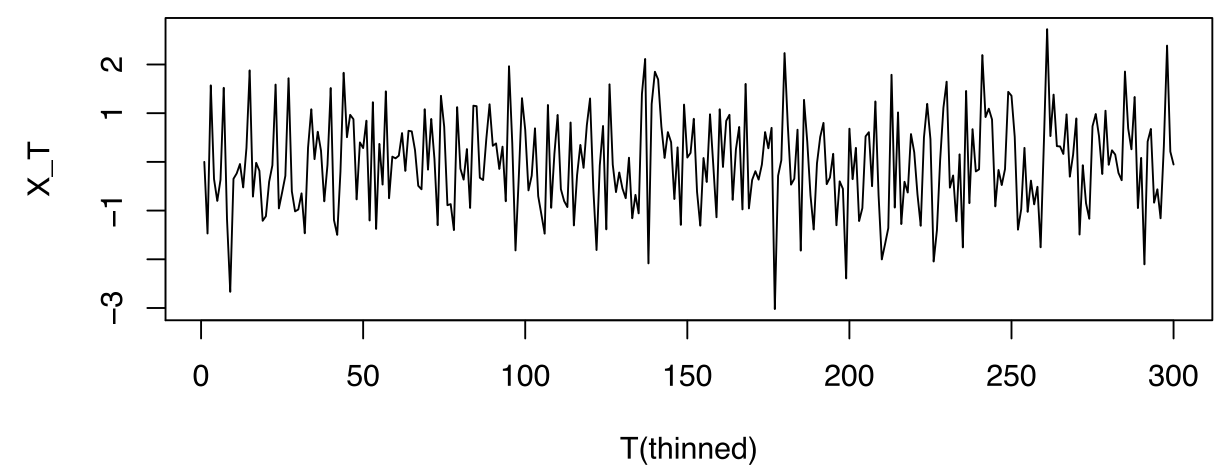
are almost independent, and have
density equal to the stationary distribution of
the Markov chain.

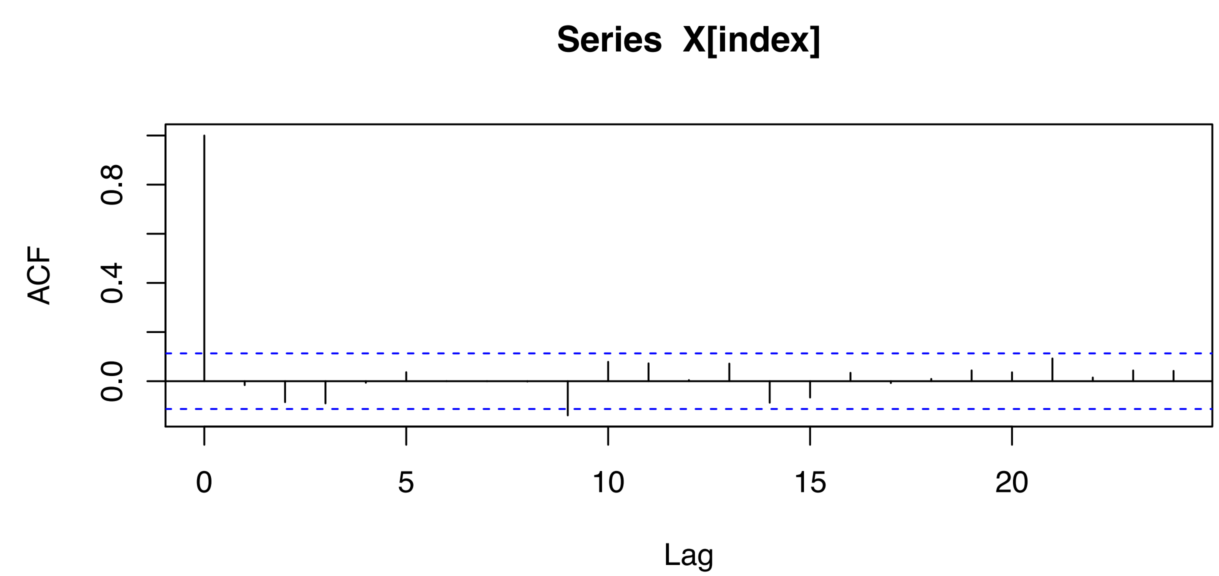
a =0.1
T=10





a =0.99
T=1000

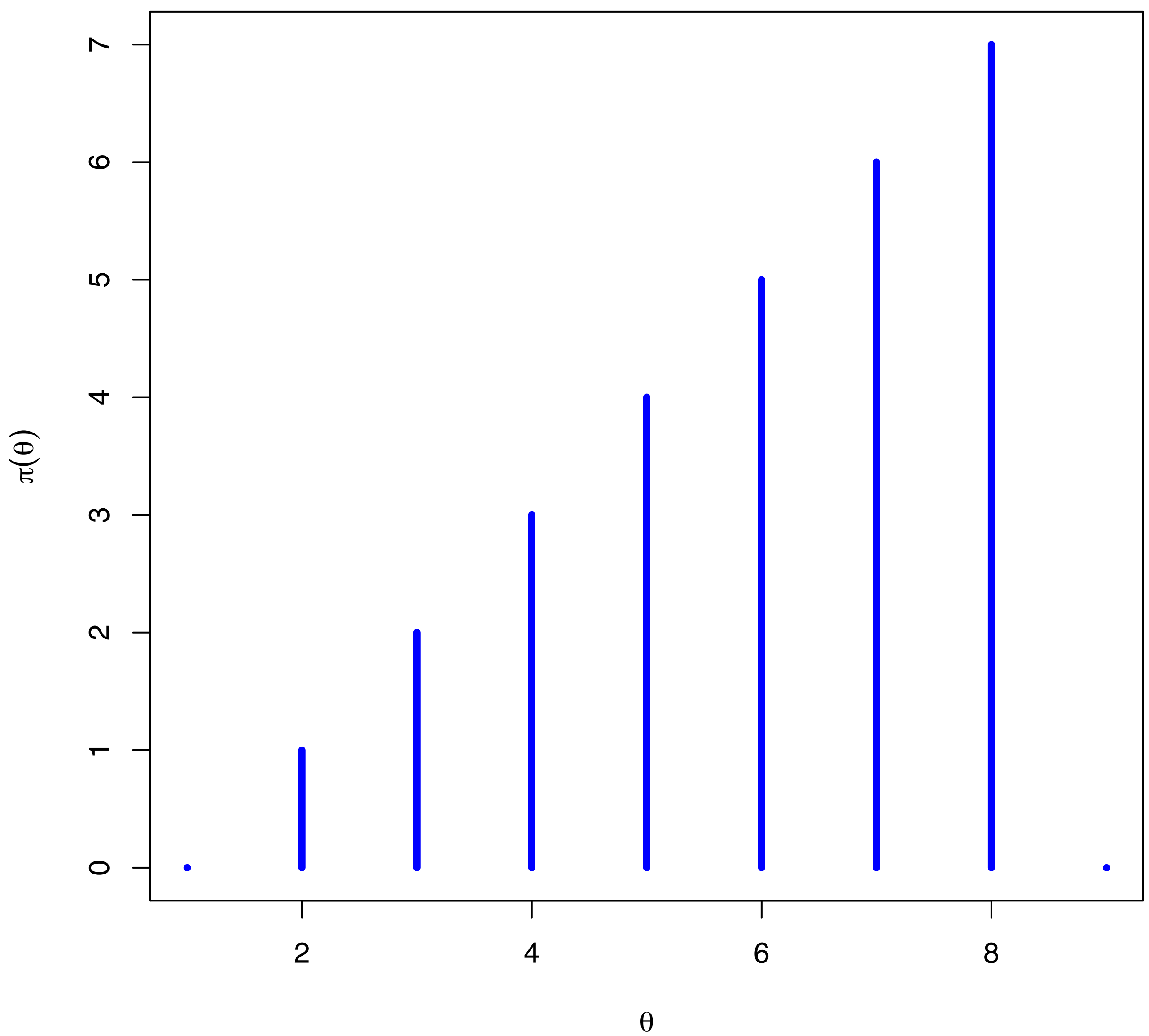




An example of a MCMC algorithm
(7.2.1 in Book)

Target unnormalized density

$$\pi(\theta = 1) = 0$$
$$\pi(\theta = 2) = 1$$
$$\pi(\theta = 3) = 2$$
$$\pi(\theta = 4) = 3$$
$$\pi(\theta = 5) = 4$$
$$\pi(\theta = 6) = 5$$
$$\pi(\theta = 7) = 6$$
$$\pi(\theta = 8) = 7$$
$$\pi(\theta = 9) = 0$$



in the book
they use p
for all densities
here use pi for
the target density
p_j for the density
of j:th iteration
of the algorithm

Proposal:

$$q(\theta_{new}|\theta_t) = \begin{cases} \theta + 1 & \text{with probability } 0.5 \\ \theta - 1 & \text{with probability } 0.5 \end{cases}$$

acceptance step:

$$U \sim U[0,1] \text{ (Beta(1,1))}$$

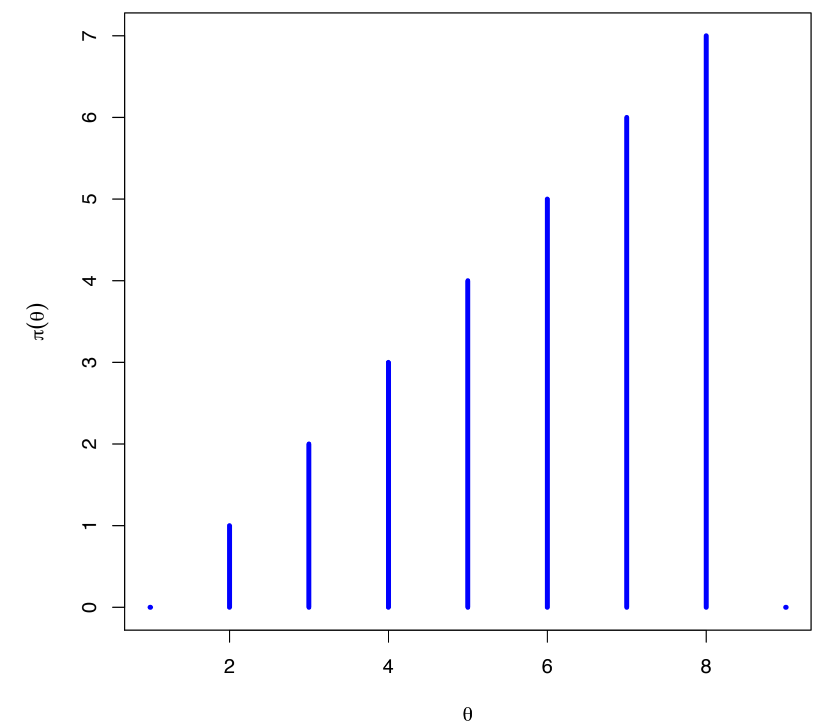
$$\text{if } U \leq \frac{p(\theta_{new})}{p(\theta_t)}$$

$$\theta_{t+1} = \theta_{new}$$

else

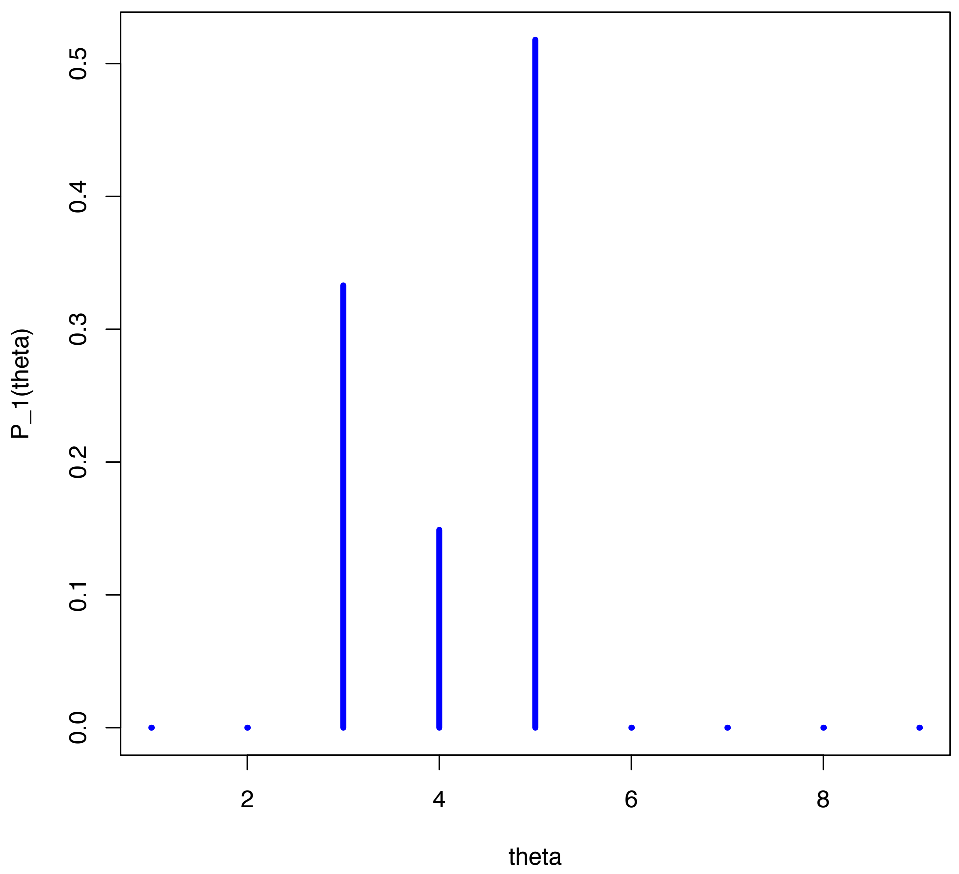
$$\theta_{t+1} = \theta_t$$

target



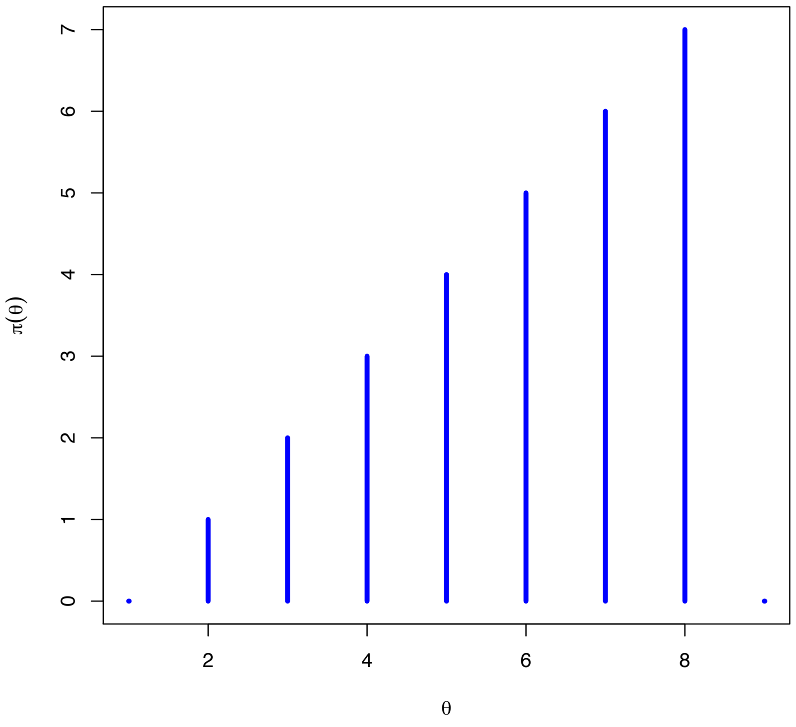
Lets examine the distribution of θ_{T+1}

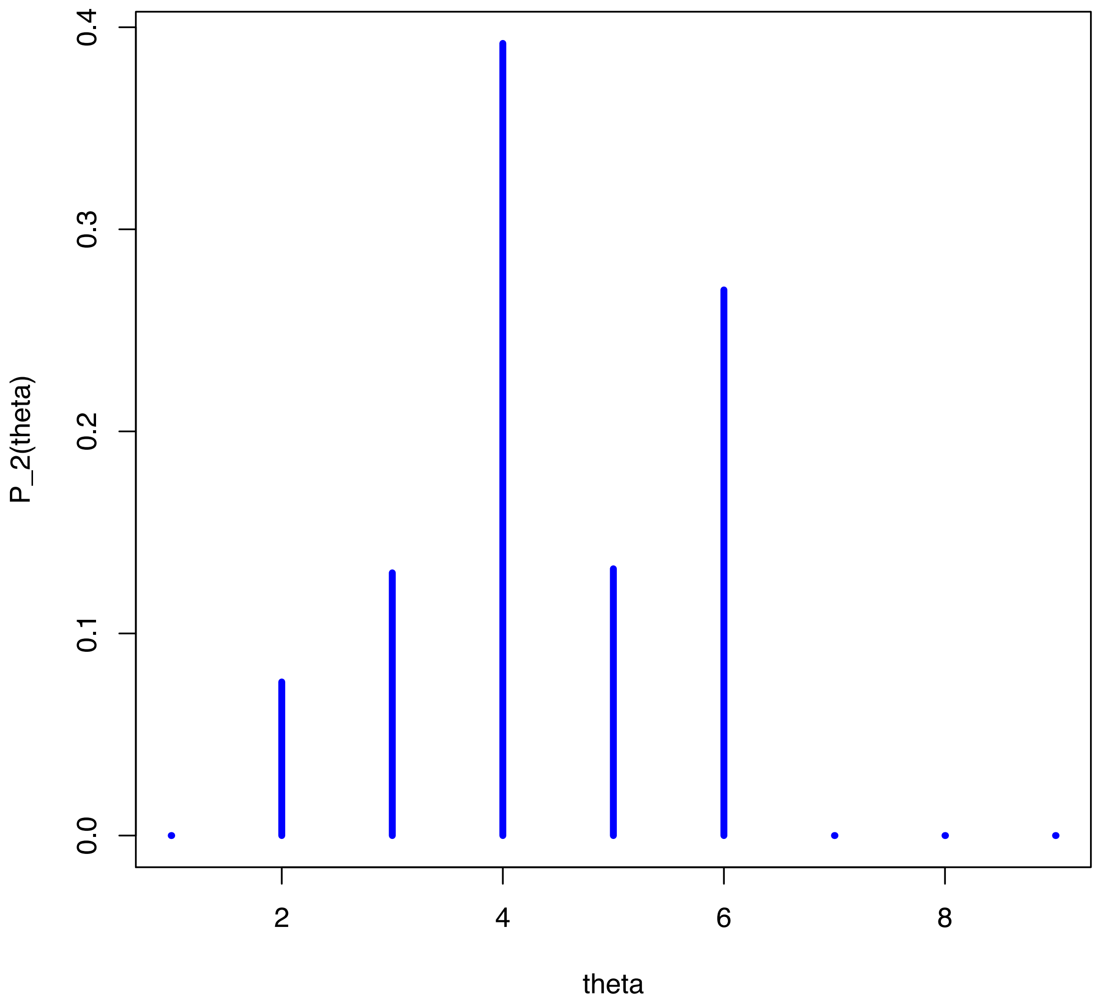
Given that $\theta_1 = 4$ for T $= 1$



Lets examine the distribution of θ_{T+1}

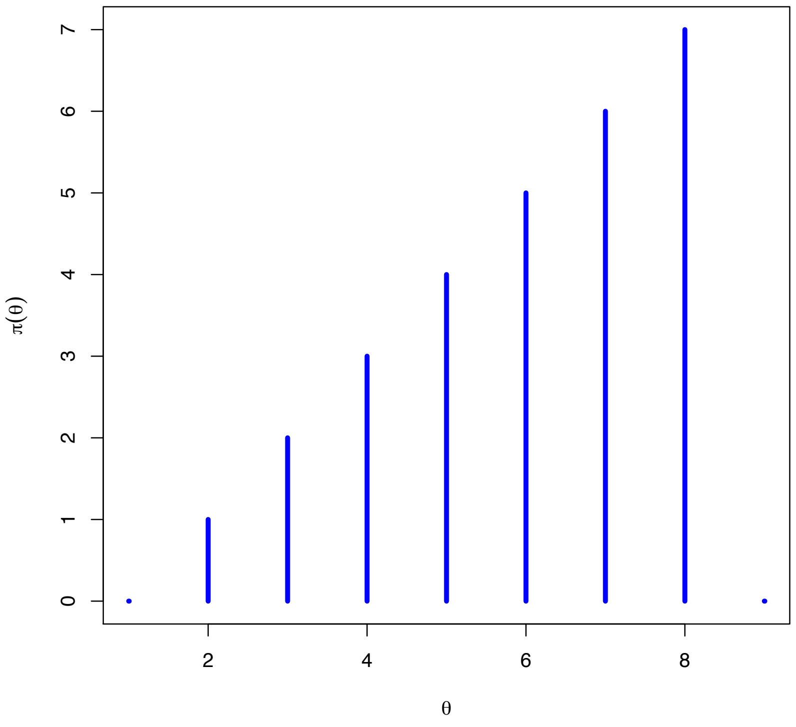
Given that $\theta_1 = 4$ for T $= 2$

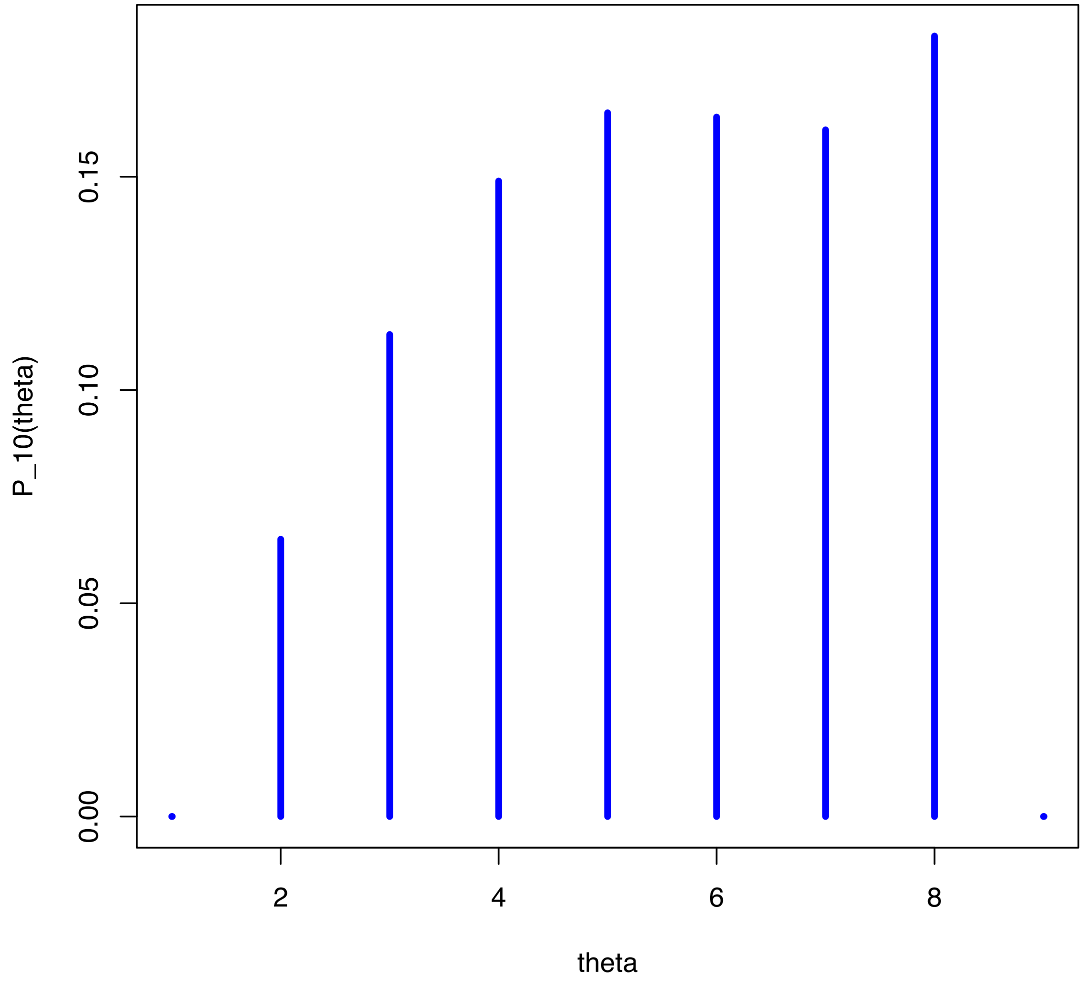




Lets examine the distribution of θ_{T+1}

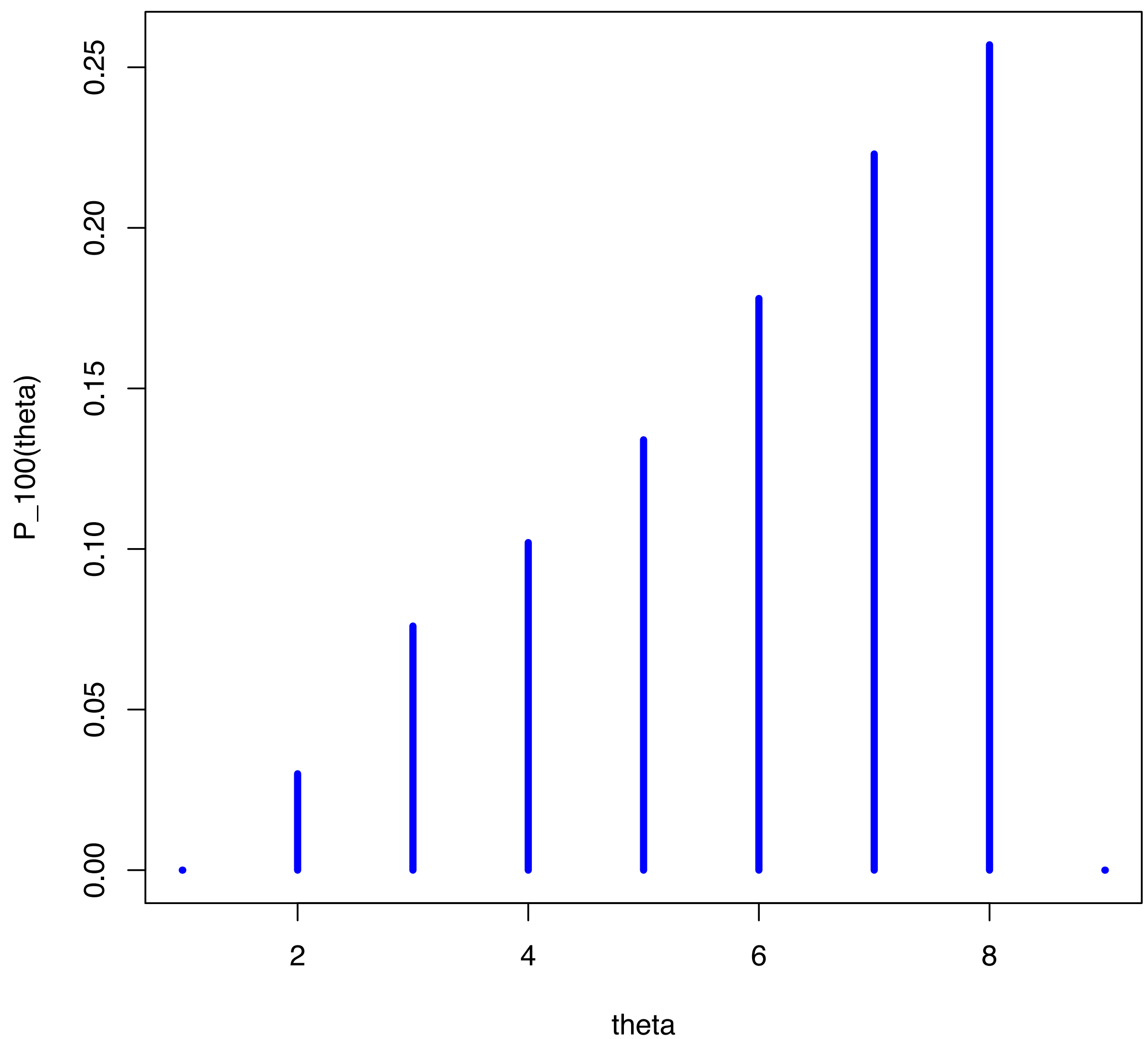
Given that $\theta_1 = 4$ for T = 10

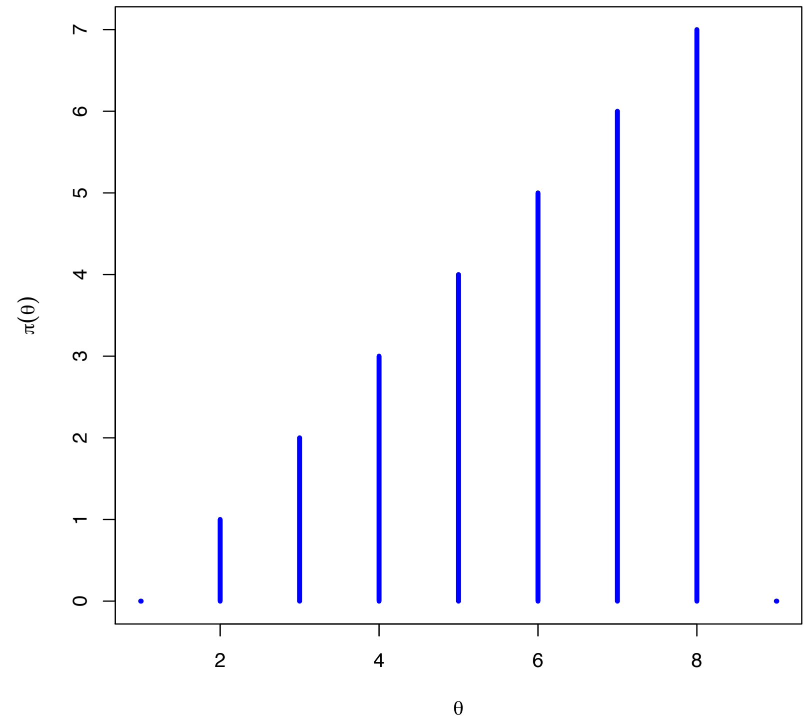




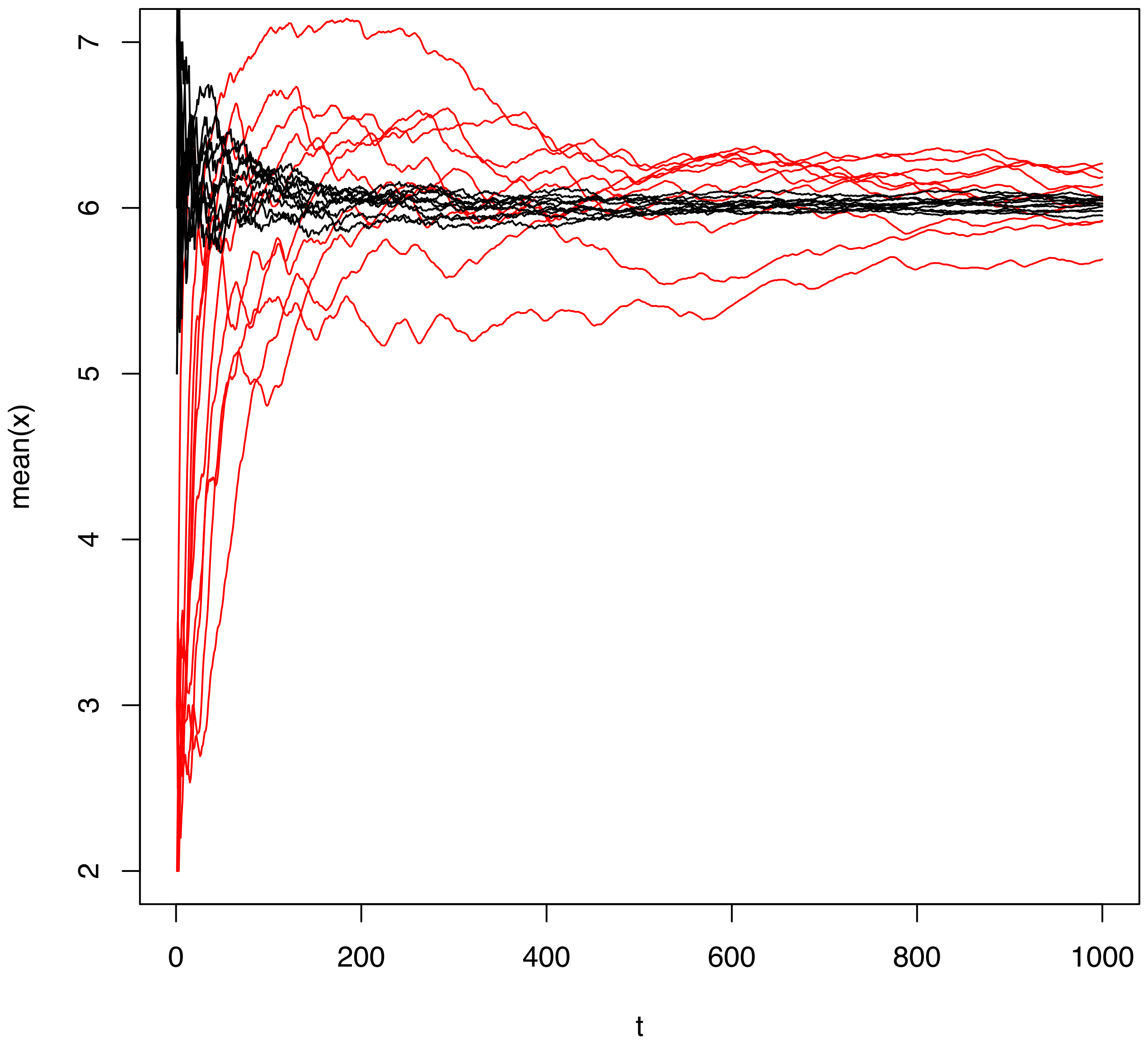
Lets examine the distribution of θ_{T+1}

Given that $\theta_1 = 4$ for T = 100





MC vs MCMC



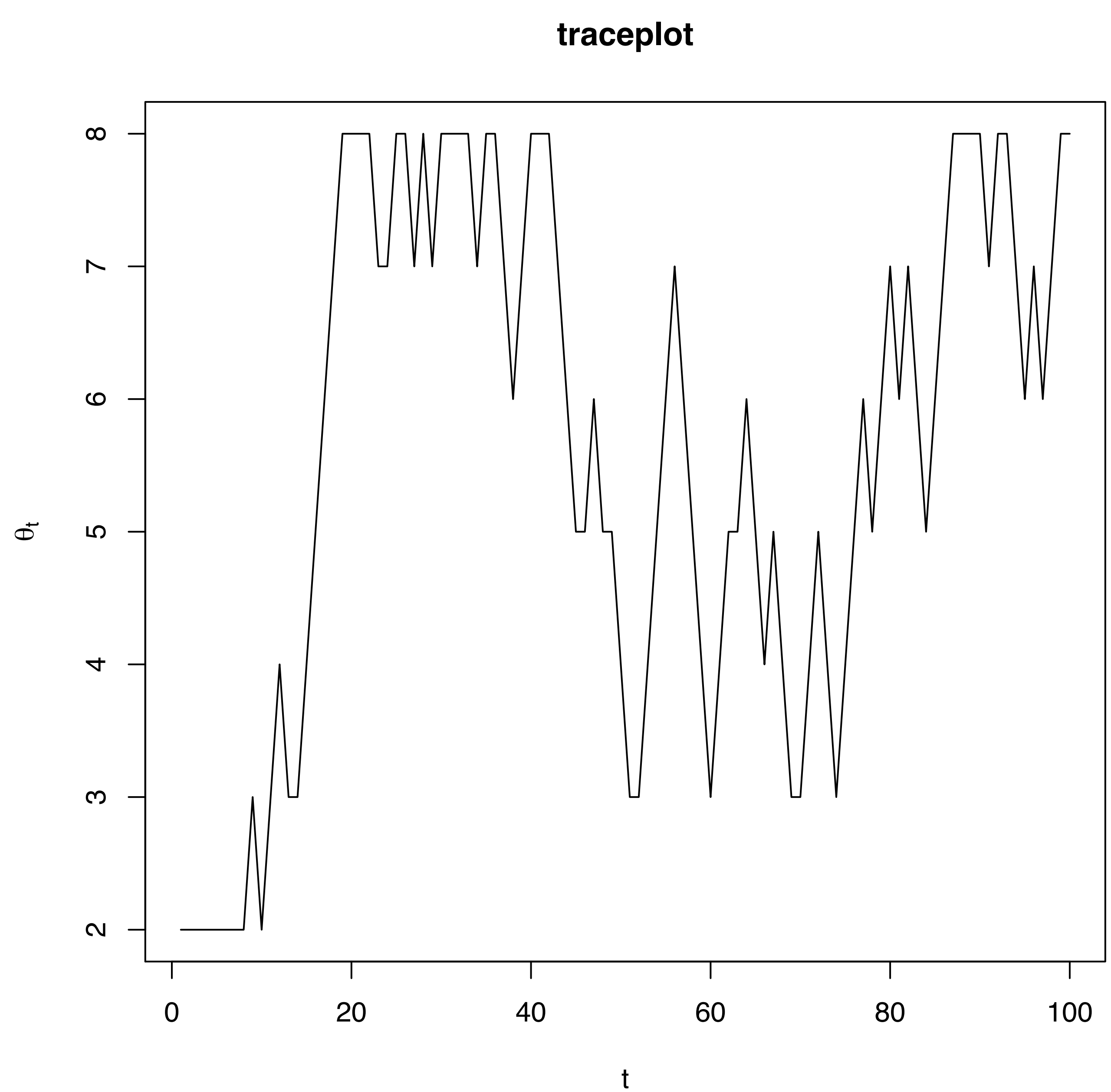
When the density is continuous
on need to use a continuous proposal like:

7.3.1 in book

$$q(\theta^{new}|\theta) = N(\theta^{new}, \theta, \sigma^2)$$

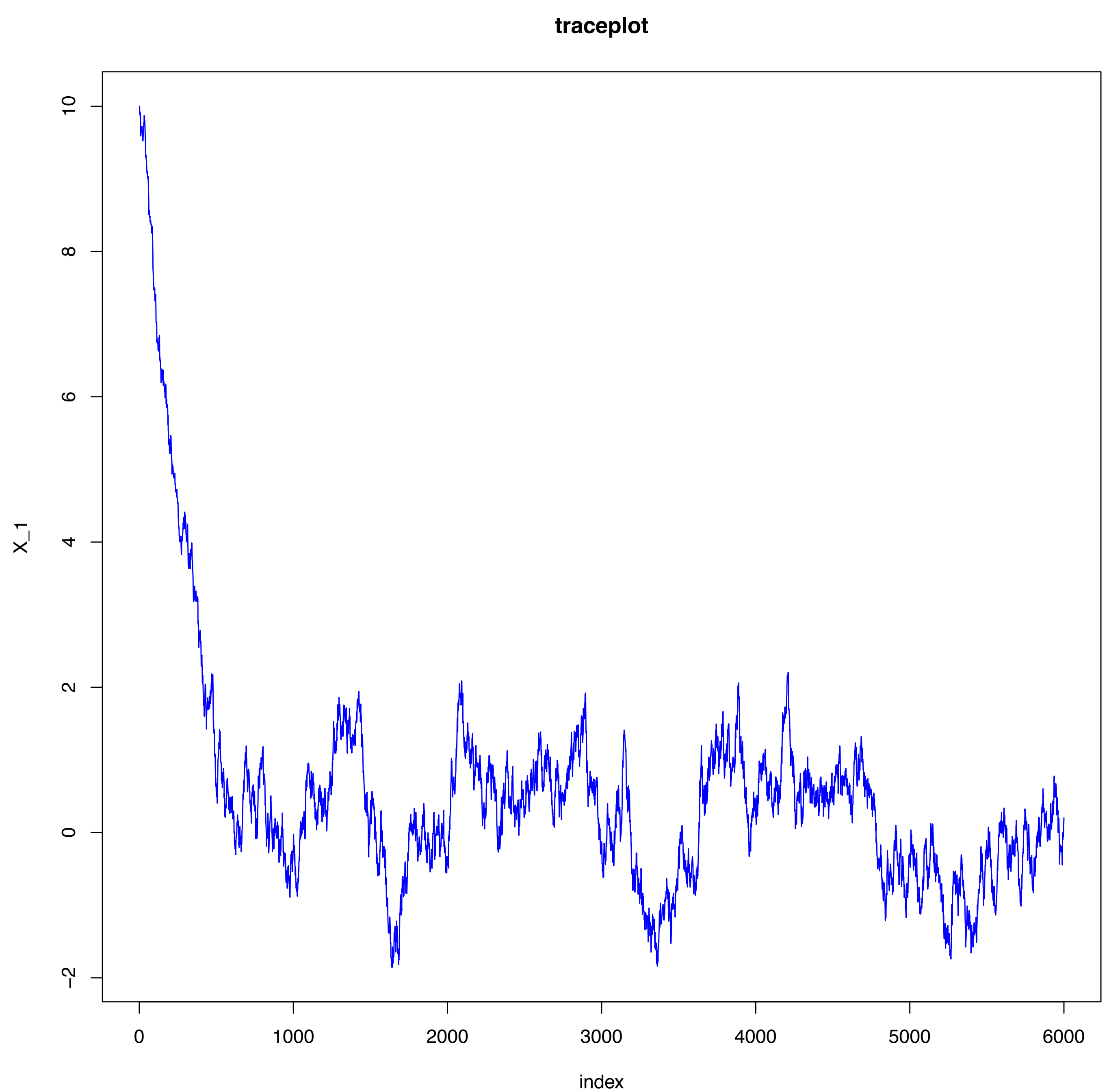
Diagnostics

Since the MCMC samples are dependent one must check so that chain has converged and that one uses enough samples

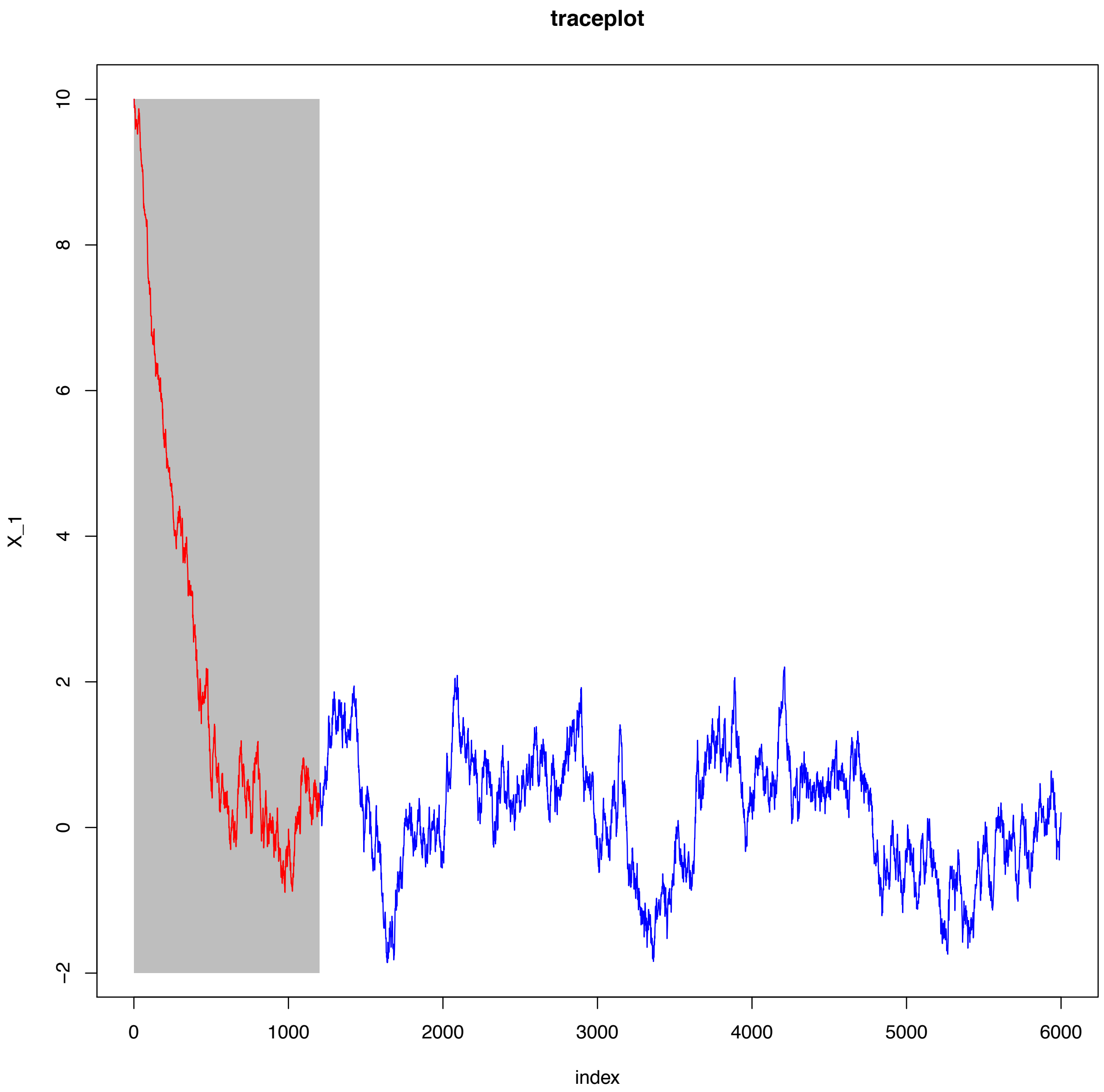


traceplot
θ_t
t
0
20
40
60
80
100
2
3
4
5
6
7
8

clear burnin:

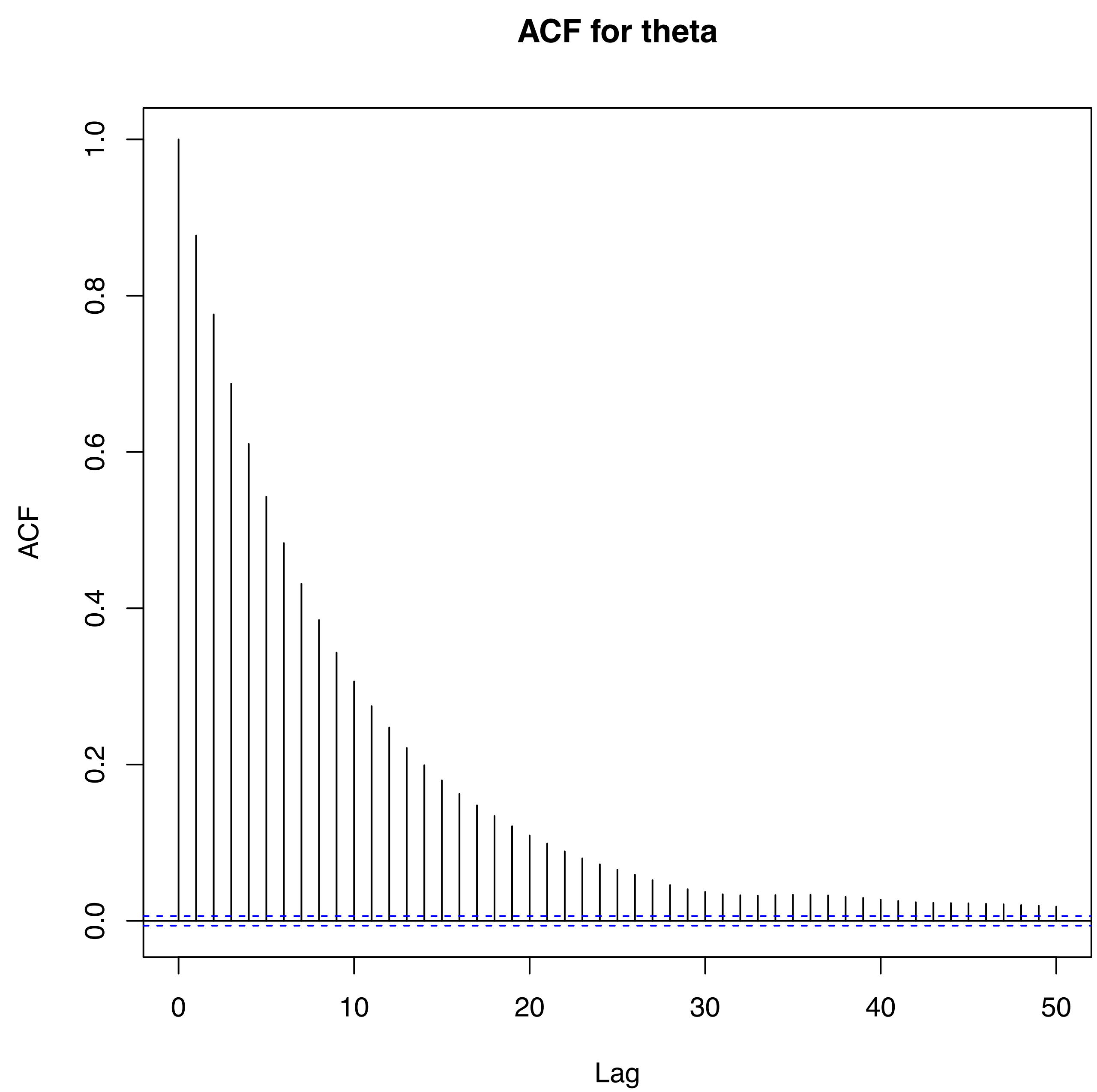


remove first part:



In stan:

$$N_{eff} = \frac{N}{1 + \sum_{k=1}^{\infty} ACF(k)}$$



A good practice is to run multiple chains and see that they converge towards the same thing: