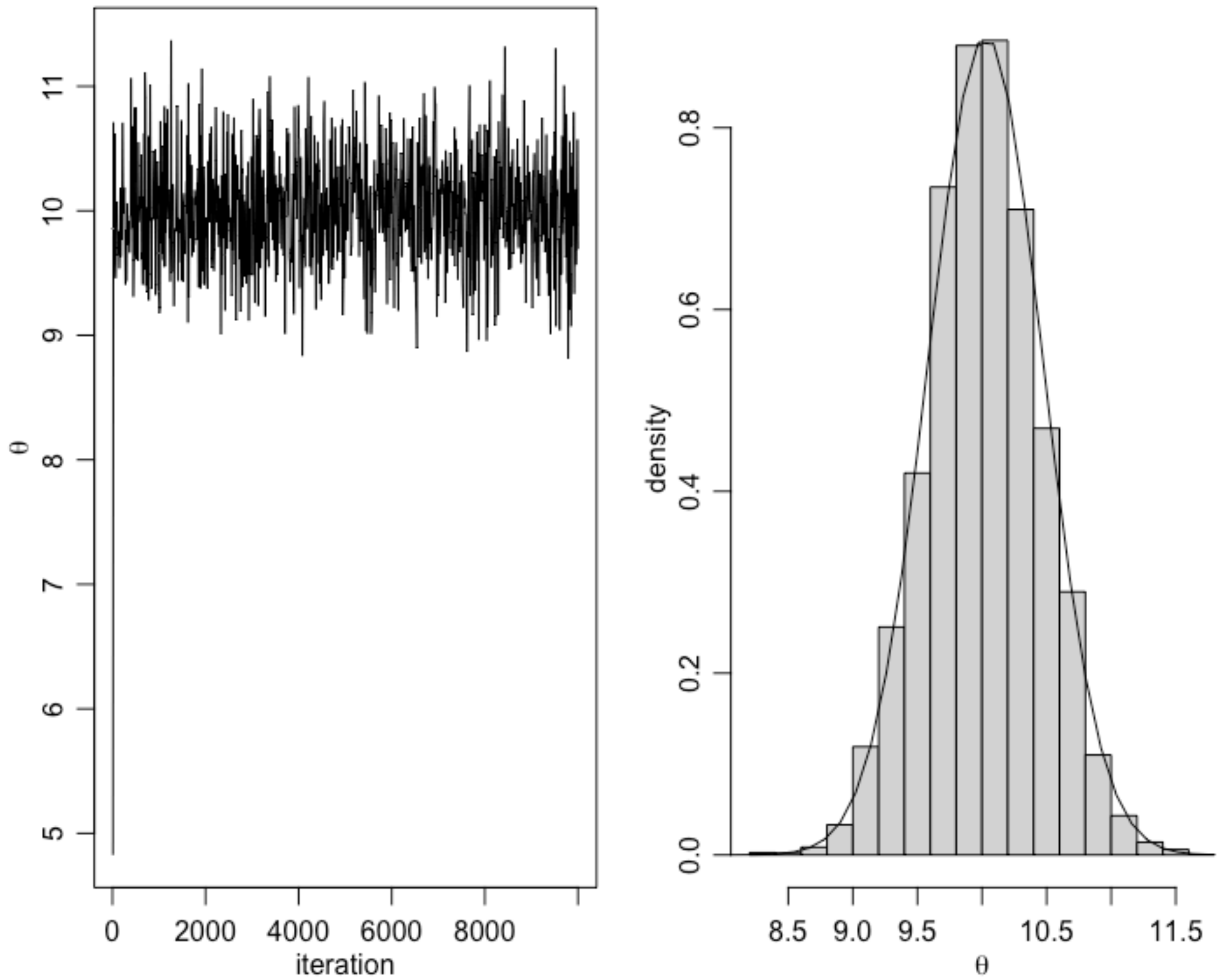
Bayes HW

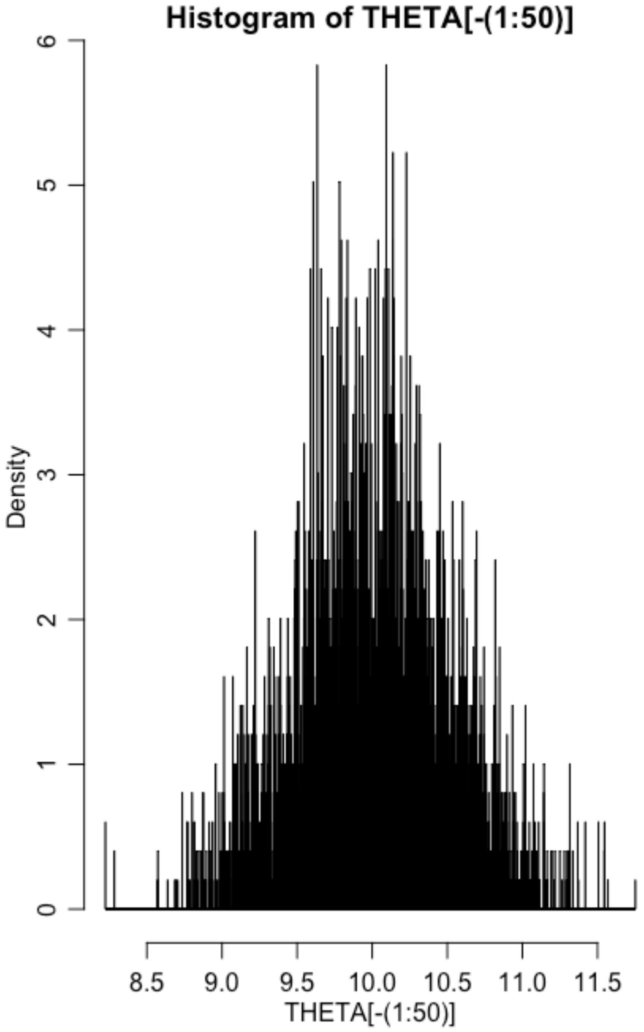
1. MCMC

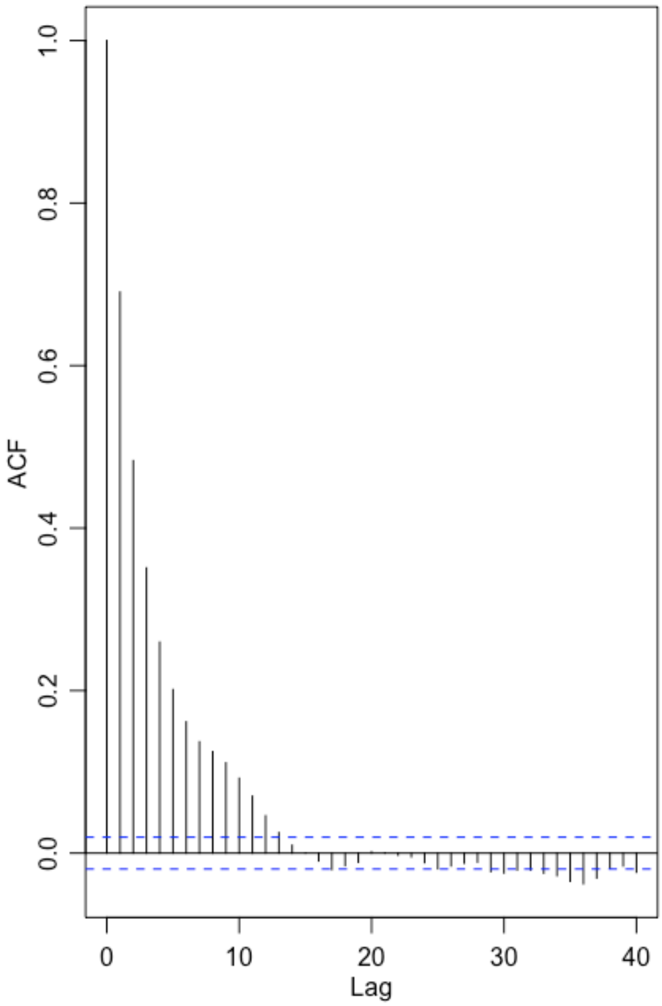
```
s2<-1 ; t2<-10 ; mu<-5 ; n<-5
y<-c(9.37, 10.18, 9.16, 11.60, 10.33)
theta<-0 ; delta<-2 ; S<-10000 ; THETA<-NULL ; set.seed(1)

mu.n<-( mean(y)*n/s2 + mu/t2 )/( n/s2+1/t2) # 직접 적분을 통해 구한 posterior
mu.n, t2.n
t2.n<-1/(n/s2+1/t2)

# MCMC
for(s in 1:S)
{
  theta.star<-rnorm(1,theta,sqrt(delta))

  log.r<-( sum(dnorm(y,theta.star,sqrt(s2),log=TRUE)) +  # log 취해서 0 근
처 값 조절
             dnorm(theta.star,mu,sqrt(t2),log=TRUE) )  -
    ( sum(dnorm(y,theta,sqrt(s2),log=TRUE)) +
        dnorm(theta,mu,sqrt(t2),log=TRUE) )

  if(log(runif(1))<log.r) { theta<-theta.star }

  THETA<-c(THETA,theta)
}

# pdf("fig10_3.pdf",family="Times",height=3.5,width=7)
par(mar=c(3,3,1,1),mgp=c(1.75,.75,0))
par(mfrow=c(1,2))

skeep<-seq(10,S,by=10)
plot(skeep,THETA[skeep],type="l",xlab="iteration",ylab=expression(theta)
)

hist(THETA[-
(1:50)],prob=TRUE,main="",xlab=expression(theta),ylab="density")
th<-seq(min(THETA),max(THETA),length=100)
lines(th,dnorm(th,mu.n,sqrt(t2.n)) )
```



- Toy example인 만큼 잘 근사가 되고 원래 분포와도 매우 비슷하게 나오는 것을 확인할 수 있음.
- 실제로는 잘 근사가 됐는지 오른쪽 히스토그램을 통해서는 알지 못할 것

```
hist(THETA[-(1:50)], breaks=9950, prob=TRUE )
acf(THETA, main="ACF")
```


Histogram of THETA[-(1:50)]
Density
0
1
2
3
4
5
6
8.5
9.0
9.5
10.0
10.5
11.0
11.5
THETA[-(1:50)]


ACF
0.0
0.2
0.4
0.6
0.8
1.0
0
10
20
30
40
Lag

2. 증명

2. HW

$\boxed{g\text{-prior : Higher } g = \text{Weaker prior}}$

$\beta_0 = 0 \quad \Sigma_0 = g\sigma^2(X^TX)^{-1}$ for any positive value g

$$\beta_n = \left(\Sigma_0^{-1} + \frac{X^TX}{\sigma^2}\right)^{-1}\left(\Sigma_0^{-1}\cdot\beta_0 + \frac{X^Ty}{\sigma^2}\right) = \left(\frac{X^TX}{g\sigma^2} + \frac{X^TX}{\sigma^2}\right)^{-1}\left(\frac{X^Ty}{\sigma^2}\right)$$

$$= \frac{g}{g+1}(X^TX)^{-1}\cancel{\sigma^2}\cdot\frac{X^Ty}{\cancel{\sigma^2}} = \underline{\frac{g}{g+1}(X^TX)^{-1}X^Ty} = \frac{g}{g+1}\hat\beta_{MLE}$$

$$\Sigma_n = \left(\Sigma_0^{-1} + \frac{X^TX}{\sigma^2}\right)^{-1} = \left(\frac{X^TX}{g\sigma^2} + \frac{X^TX}{\sigma^2}\right)^{-1} = \underline{\frac{g}{g+1}(X^TX)^{-1}\sigma^2} = \frac{g}{g+1}V(\hat\beta_{MLE})$$

$$\Rightarrow \beta \mid y, \sigma^2 \sim N\left(\frac{g}{g+1}\hat\beta_{MLE}, \frac{g}{g+1}V(\hat\beta_{MLE})\right)$$

$SSR(g) \longrightarrow SSR(\hat\beta_{MLE})$? $\qquad m = \beta_n$

$$SSR(g) = y^Ty - \sigma^2 m^TV^{-1}m = y^Ty - \cancel{\sigma^2}\boxed{\frac{\cancel{g}}{\cancel{g+1}}y^TX\cdot\cancel{(X^TX)^{-1}}}\cdot\boxed{\frac{\cancel{g+1}}{\cancel{g}}\frac{1}{\cancel{\sigma^2}}\cancel{(X^TX)}}\cdot\boxed{\frac{g}{g+1}(X^TX)^{-1}X^Ty}$$

$$= y^T\left(I - \frac{g}{g+1}X(X^TX)^{-1}X^T\right)y$$

$$\approx y^T\left(I - X(X^TX)^{-1}X^T\right)y \quad \text{as } g\to\infty$$

$$SSR(\hat\beta_{MLE}) = (y - X\hat\beta_{MLE})^T(y - X\hat\beta_{MLE})$$

$$= (y - X(X^TX)^{-1}X^Ty)^T(y - X(X^TX)^{-1}X^Ty)$$

$$= (y^T - y^TX(X^TX)^{-1}X^T)(y - X(X^TX)^{-1}X^Ty)$$

$$= y^Ty - 2\cdot y^TX(X^TX)^{-1}X^Ty + y^TX(X^TX)^{-1}\cancel{X^TX(X^TX)^{-1}}X^Ty$$

$$= y^Ty - y^TX(X^TX)^{-1}X^Ty$$

$\therefore\ SSR(g) \approx SSR(\hat\beta_{MLE})$ as $g\to\infty$