Hierarchical modeling: The files `school1.dat` through `school8.dat` give weekly hours spent on homework for students sampled from eight different schools. Obtain posterior distributions for the true means for the eight different schools using a hierarchical normal model with the following prior parameters:

$$\mu_0 = 7, \gamma_0^2 = 5, \quad \tau_0^2 = 10, \eta_0 = 2, \quad \sigma_0^2 = 15, \nu_0 = 2\,.$$

a) Run a Gibbs sampling algorithm to approximate the posterior distribution of $\{\boldsymbol{\theta}, \sigma^2, \mu, \tau^2\}$. Assess the convergence of the Markov chain, and find the effective sample size for $\{\sigma^2, \mu, \tau^2\}$. Run the chain long enough so that the effective sample sizes are all above 1,000.

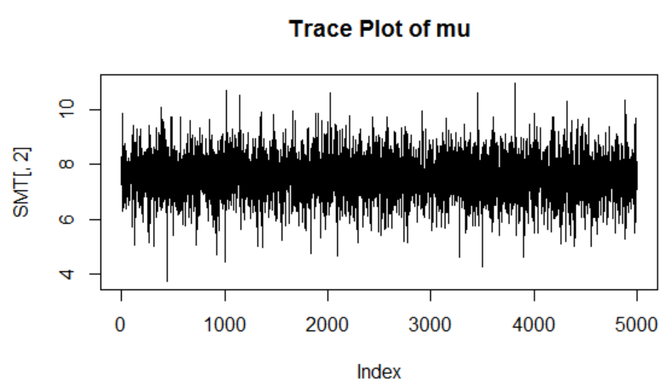
b) Compute posterior means and 95% confidence regions for $\{\sigma^2, \mu, \tau^2\}$. Also, compare the posterior densities to the prior densities, and discuss what was learned from the data.

c) Plot the posterior density of $R = \frac{\tau^2}{\sigma^2+\tau^2}$ and compare it to a plot of the prior density of R. Describe the evidence for between-school variation.

d) Obtain the posterior probability that θ_7 is smaller than θ_6, as well as the posterior probability that θ_7 is the smallest of all the θ's.

e) Plot the sample averages $\bar{y}_1, \ldots, \bar{y}_8$ against the posterior expectations of $\theta_1, \ldots, \theta_8$, and describe the relationship. Also compute the sample mean of all observations and compare it to the posterior mean of μ.

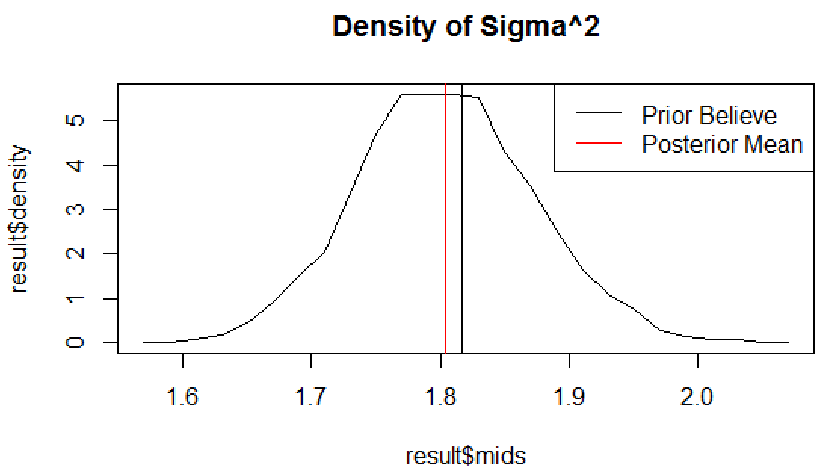
a) Run Gibbs sampling and approximate the posterior distribution. The code is copied in attachment. Results are stored in variable "THETA" and "SMT". Check convergence, we know the result is acceptable. Use post-burn-in values for following calculations.

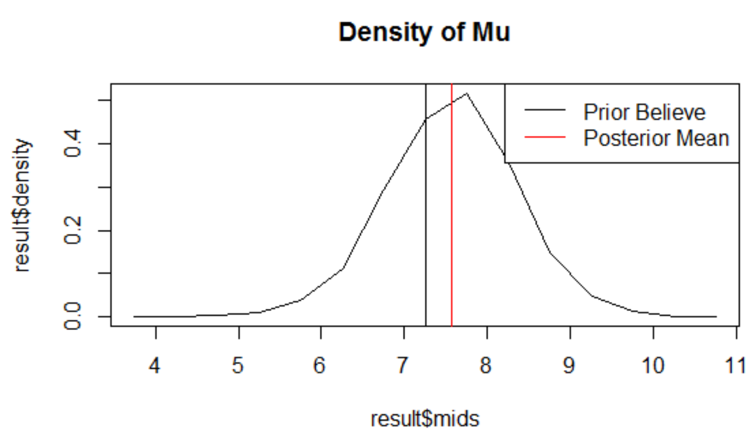


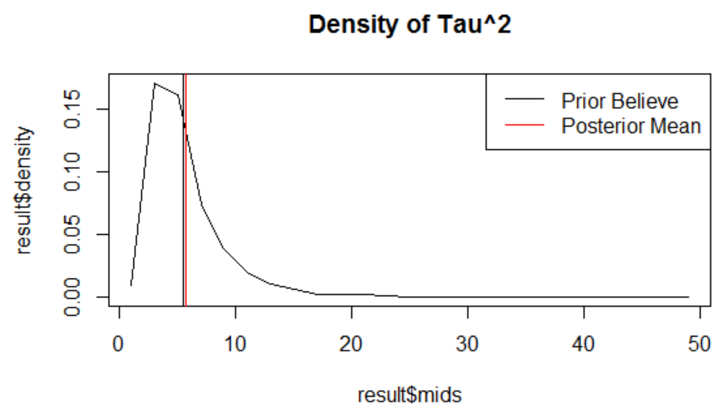
b)

	Prior	Posterior Mean	95% Confidence Interval
σ^2	1.816596	1.803604	(1.672684, 1.941098)
μ	7.262718	7.567186	(5.983655, 9.081597)
τ^2	5.445397	5.661819	(2.077722, 14.510708)

Comparisons between the densities are drawn out below:

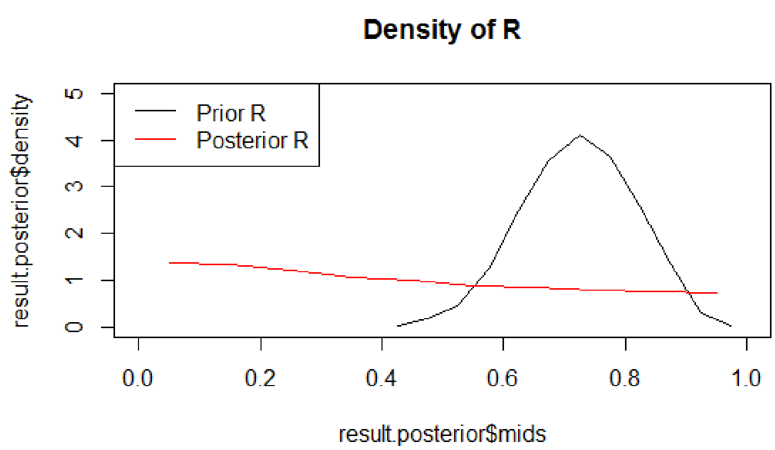






Comparing the densities we can see that the posterior mean is not very far away from our prior belief about the data. The sampling gave us more confident of the area where those variables may fall in.

3) The posterior density of R is plotted below with comparison of prior.



	Mean	95% Confidence Interval
Prior R	0.4350736	(0.0178848, 0.9631193)
Posterior R	0.7221168	(0.5347478, 0.8896042)

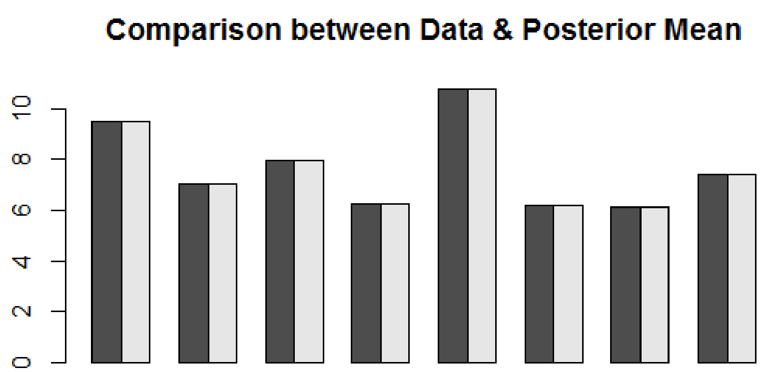
The posterior belief in R shows that “between school variance tau^2” is actually bigger and more dominant than the “within school variance sigma^2”, which resulted in that over 95% R >0.5.

d) The probability is calculated using the result from Gibbs sampling, whose code is attached.

	Probability
$\theta 7 < \theta 6$	65.56%
$\theta 7 <$ All other θ	54.12%

e) Data and posterior mean is shown in the below table and plotted as well. We can tell that the difference between data mean and posterior mean is pretty small, as a result of a weakly informative prior.

School	1	2	3	4	5	6	7	8
Data	9.464000	7.033478	7.953000	6.232083	10.765833	6.205000	6.132727	7.381000
Posterior	9.462212	7.033666	7.950210	6.232945	10.761551	6.207202	6.140305	7.380687



The comparison between all sample data mean and posterior mean of mu is shown in the below table, we can see that the posterior gives lower mean than the data. The reason for this is because in the data, there are more schools lower than the average and they dominate the posterior estimate of mu.

All Data Mean	7.646097
Posterior Mean of mu	7.262718

```
###CODE

library(mvtnorm)
#data
m=8
Y<-matrix(ncol=m)
Y[1] <- read.table("http://www.stat.washington.edu/hoff/Book/Data/hwdata/school1.dat")
Y[2] <- read.table("http://www.stat.washington.edu/hoff/Book/Data/hwdata/school2.dat")
Y[3] <- read.table("http://www.stat.washington.edu/hoff/Book/Data/hwdata/school3.dat")
Y[4] <- read.table("http://www.stat.washington.edu/hoff/Book/Data/hwdata/school4.dat")
Y[5] <- read.table("http://www.stat.washington.edu/hoff/Book/Data/hwdata/school5.dat")
Y[6] <- read.table("http://www.stat.washington.edu/hoff/Book/Data/hwdata/school6.dat")
Y[7] <- read.table("http://www.stat.washington.edu/hoff/Book/Data/hwdata/school7.dat")
Y[8] <- read.table("http://www.stat.washington.edu/hoff/Book/Data/hwdata/school8.dat")

#as.matrix(as.data.frame())
#X<-matrix(nrow=25,ncol=m)

##weakly informative priors
nu0<-2 ; s20<-15
eta0<-2 ; t20<-10
mu0<-7 ; g20<-5

### initials

n<-sv<-ybar<-rep(NA,m)
for ( j in 1:m)
{
  y<-unlist(Y[j])
  ybar[j] <- mean(y)
  sv[j] <- var(y)
  n[j] <- sum(y)
}
theta <- ybar ; sigma2 <- mean(sv)
mu <- mean(theta) ; tau2 <-var(theta)

#########GIBBS#######
set.seed(1)
S<-5000
THETA<-matrix(nrow=S , ncol=m)
SMT<-matrix(nrow=S , ncol=3)

###
for ( s in 1:S)
{
  #sample a new value of mu
  vmu<- 1/(m/tau2+1/g20 )
  emu<- vmu*(m*mean(theta)/tau2 + mu0/g20 )
```

```
  mu<-rnorm( 1,emu,sqrt(vmu))


  #sample a new value of tau2
  etam<- eta0+m
  ss<- eta0*t20 + sum( (theta-mu)^2 )
  tau2<- 1/rgamma ( 1 , etam/2 , ss/2)

  #sample new values of the thetas
  for (j in 1:m)
  {
   vtheta <- 1/(n[j]/sigma2+1/tau2 )
   etheta <- vtheta*( ybar[j] * n[j]/sigma2+mu/tau2 )
   theta[j] <- rnorm( 1 , etheta , sqrt(vtheta))
  }

  #sample new value of sigma2
  nun<- nu0+sum(n)
  ss<- nu0*s20
  for ( j in 1 :m)
  {
   y=unlist(Y[j])
   ss<-ss+sum((y - theta[j])^2)
  }
  sigma2 <- 1/rgamma ( 1,nun/2,ss/2)

  #store results
  THETA[s,]<-theta
  SMT[s,]<-c(sigma2,mu,tau2)
}
plot(SMT[,2],type="l",main='Trace Plot of mu')

##########sigma
mean(SMT[,1])
quantile(SMT[,1],c(0.025,0.975))
result<-hist(SMT[,1],breaks=20,plot="false")
plot(result$mids,result$density,type="l",main="Density of Sigma^2")
abline(v=sigma2,col=1)
abline(v=mean(SMT[,1]),col=2)
legend("topright", legend=c("Prior Believe","Posterior Mean"),lty=c(1,1),col=c(1,2))

###### mu
mean(SMT[,2])
quantile(SMT[,2],c(0.025,0.975))
result<-hist(SMT[,2],breaks=20,plot="false")
plot(result$mids,result$density,type="l",main="Density of Mu")
abline(v=mu,col=1)
abline(v=mean(SMT[,2]),col=2)
```

```
legend("topright", legend=c("Prior Believe","Posterior Mean"),lty=c(1,1),col=c(1,2))

###### tau2
mean(SMT[,3])
quantile(SMT[,3],c(0.025,0.975))
result<-hist(SMT[,3],breaks=20,plot="false")
plot(result$mids,result$density,type="l",main="Density of Tau^2")
abline(v=tau2,col=1)
abline(v=mean(SMT[,3]),col=2)
legend("topright", legend=c("Prior Believe","Posterior Mean"),lty=c(1,1),col=c(1,2))

# Question 3
#posterior
R=SMT[,3]/(SMT[,1]+SMT[,3])
mean(R)
quantile(R,c(0.025,0.975))
result.posterior<-hist(R,plot="false")
#prior
tau2.prior<-1/rgamma(5000,eta0/2,eta0*t20/2)
sigma2.prior<-1/rgamma(5000,nu0/2,nu0*s20/2)
R.prior=tau2.prior/(sigma2.prior+tau2.prior)
mean(R.prior)
quantile(R.prior,c(0.025,0.975))
result.prior<-hist(R.prior,plot="false")

plot(result.posterior$mids,result.posterior$density,type="l",main="Density of R",xlim=c(0,1),ylim=c(0,5))
lines(result.prior$mids,result.prior$density,type="l",col=2)
legend("topleft", legend=c("Prior R","Posterior R"),lty=c(1,1),col=c(1,2))


######### Problem 4
####4.1
count=0
for ( s in 1:S)
{
  if (THETA[s,7]<THETA[s,6]) count=count+1
}
print(count/S)

##4.2
s=1
count=0
for ( s in 1:S)
{
  count2=0
  for (j in 1:m)
  {
    if (THETA[s,7]<THETA[s,j]) count2=count2+1
```

```
  }
  if (count2==7) count=count+1
}
print(count/S)

#####Problem 5
colMeans(THETA)
ybar
mean(mu)
mean(THETA)
```