Paul Cubre

Math 9810 Homework 3

5.1 Studying: The files school1.dat, school2.dat, and school3.dat contain data on the amount of time students from three high schools spent on studying or homework during an exam period. Analyze data form each of these schools separately, using the normal model with a conjugate prior distribution, in which $\{\mu_0 = 5, \sigma_0^2 = 4, \kappa_0 = 1, \nu_0 = 2\}$ and compute or approximate the following:

a) posterior means and 95% confidence intervals for the mean θ and standard deviation σ from each school;

We may use Monte Carlo sampling of 10^6 to estimate the following:

	Mean θ	95% CI	Mean σ	95% CI
School 1	9.292256	$(7.764692, 8.301262)$	3.907923	$(3.001864, 3.248633)$
School 2	6.947922	$(5.159098, 5.787117)$	4.397937	$(3.345090, 3.630897)$
School 3	7.813435	$(6.176921, 6.758262)$	3.748485	$(2.800274, 3.052971)$

Where $\sigma^{(s)} \sim Inv - Gamma(\nu_n/2, \sigma_n^2\nu_n/2)$ and $\theta^{(s)} \sim norm(\mu_n, \sigma^{2(s)}/\kappa_n)$.

b) The posterior probability that $\theta_i < \theta_j < \theta_k$ for all six permutations $\{i, j, k\}$ of $\{1, 2, 3\}$;

We may use the data from the Monte Carlo to estimate the probabilities. Therefore we have $P(\theta_i < \theta_j < \theta_k) = \{0.005869, 0.003838, 0.084452, 0.672709, 0.015373, 0.217759\}$ when $\{i, j, k\}$ are $\{123, 132, 213, 231, 312, 321\}$ respectively.

c) The posterior probability that $\bar{Y}_i < \bar{Y}_j < \bar{Y}_k$ for all six permutations $\{i, j, k\}$ of $\{1, 2, 3\}$, where $\bar{Y}_i$ is a sample from the posterior predictive distribution of school i;

We may use the data from the Monte Carlo to estimate the probabilities. We have that $\bar{Y}^{(s)} \sim norm(\theta^{(s)}, \sigma^{2(s)}/\kappa_n)$. Therefore we have $P(\bar{Y}_i < \bar{Y}_j < \bar{Y}_k) = \{0.021333, 0.015134, 0.138122, 0.534346, 0.0444$ when $\{i, j, k\}$ are $\{123, 132, 213, 231, 312, 321\}$ respectively.

d) Compute the posterior probability that θ_1 is bigger than both θ_2 and θ_3, and the posterior probability that $\bar{Y}_1$ is bigger than both $\bar{Y}_2$ and $\bar{Y}_3$.

We may simply take $P(\theta_1 > \theta_2, \theta_3) = P(\theta_3 < \theta_2 < \theta_1) + P(\theta_2 < \theta_3 < \theta_1) = 0.780936$ and $P(\bar{Y}_1 > \bar{Y}_2, \bar{Y}_3) = P(\bar{Y}_3 < \bar{Y}_2 < \bar{Y}_1) + P(\bar{Y}_2 < \bar{Y}_3 < \bar{Y}_1) = 0.890468$.

6.1 Poisson population comparisons: Let's reconsider the number of children data of Exercise 4.8. We'll assume Poisson sampling models for the two groups as before, but now we'll parameterize θ_A and θ_B as $\theta_A = \theta$ and $\theta_B = \theta \times \gamma$. In this parameterization, γ represents the relative rate θ_B/θ_A. Let $\theta \sim gamma(a_\theta, b_\theta)$ and let $\gamma \sim gamma(a_\gamma, b_\gamma)$.

a) Are θ_A and θ_B independent or dependent under this prior distribution? In what situations is such a joint prior distribution justified?

We have that $P(\theta_B|\theta_A) = P(\gamma) \neq P(\theta_B)$. Therefore we have a strong dependence unless γ is evenly distributed.

b) Obtain the form of the full conditional distribution of θ given y_A, y_B and γ.
Let Y_A be X and Y_B be Y. Then

$$
\begin{aligned}
P(\theta,\gamma|x,y) &\propto P(x,y|\gamma,\theta)P(\gamma)P(\theta) \\
&= \prod P(x|\theta)\prod P(y|\theta,\gamma)P(\gamma)P(\theta) \\
&= e^{-n_A\theta}\frac{\theta^{\sum x_i}}{\prod x_i}e^{-n_B\theta\gamma}\frac{(\theta\gamma)^{\sum y_i}}{\prod y_i}\frac{b_\theta^{a_\theta}}{\Gamma(a_\theta)}\theta^{a_\theta-1}e^{-b_\theta\theta}\frac{b_\gamma^{a_\gamma}}{\Gamma(a_\gamma)}\gamma^{a_\gamma-1}e^{-b_\gamma\gamma} \\
&= P(\theta|\gamma,x,y)P(\gamma|x,y) \\
&= P(\theta|x,y)P(\gamma|x,y,\theta)
\end{aligned}
$$

Thus we isolate θ to get the conditional probability

$$P(\theta|\gamma,x,y) \propto e^{-(n_A+n_B\gamma+b_\theta)\theta}\theta^{\sum x_i+\sum y_i+a_\theta-1}$$

Thus $\theta|\gamma, y_A, y_B \sim gamma(\sum y_A + \sum y_B + a_\theta, b_\theta + n_A + \gamma n_B)$.

c) Obtain the form of the full conditional distribution of γ given y_A, y_B and θ.
Again we isolate the γ to get the conditional probability

$$P(\gamma|\theta,x,y) \propto e^{-(n_B\theta+b_\gamma)\gamma}\gamma^{\sum y_i+a_\gamma-1}$$

Thus $\gamma|\theta, y_A, y_B \sim gamma(\sum y_B + a_\gamma, b_\gamma + n_B\theta)$.

d) Set $a_\theta = 2$ and $b_\theta = 1$. Let $a_\gamma = b_\gamma \in \{8, 16, 32, 64, 128\}$. For each of these five values, run a Gibbs sampler of at least 5,000 iterations and obtain $E[\theta_B - \theta_A|y_A, y_B]$. Describe the effects of the prior distribution for γ on the results.

We have that Gibbs sampler computed $E[\theta_B-\theta_A|y_A, y_B] = \{0.3851789, 0.3340760, 0.2700603, 0.1998369, 0.1$ for $\gamma = \{2^3, 2^4, 2^5, 2^6, 2^8\}$ which is fairly linear. Thus we expect when one doubles $a_\gamma = b_\gamma$ that the mean difference between the posterior θ's decrease linearly.

6.2 Mixture model: The file glucose.dat contains the plasma glucose concentration of 532 females from a study on diabetes (see Exercise 7.6).

a) Make a histogram or kernel density estimate of the data. Describe how this empirical distribution deviates from the shape of a normal distribution.
In Figure 1, we see that the distribution is skewed to the right. Hence we do not have a nice bell shaped and symmetric density.

b) Consider the following mixture model for these data: For each study participant there is an unobserved group membership variable X_i which is equal to 1 or 2 with probability p and $1-p$. If $X_i = 1$ then $Y_i \sim normal(\theta_1, \sigma_1^2)$, and if $X_i = 2$ then $Y_i \sim normal(\theta_2, \sigma_2^2)$. Let $p \sim beta(a, b)$, $\theta_j \sim normal(\mu_0, \tau_0^2)$ and $1/\sigma_j^2 \sim gamma(\nu_0/2, \nu_0\sigma_0^2/2)$ for both $j = 1$ and $j = 2$. Obtain the full conditional distributions of $(X_1, \ldots, X_n), p, \theta_1, \theta_2, \sigma_1^2$, and σ_2^2.

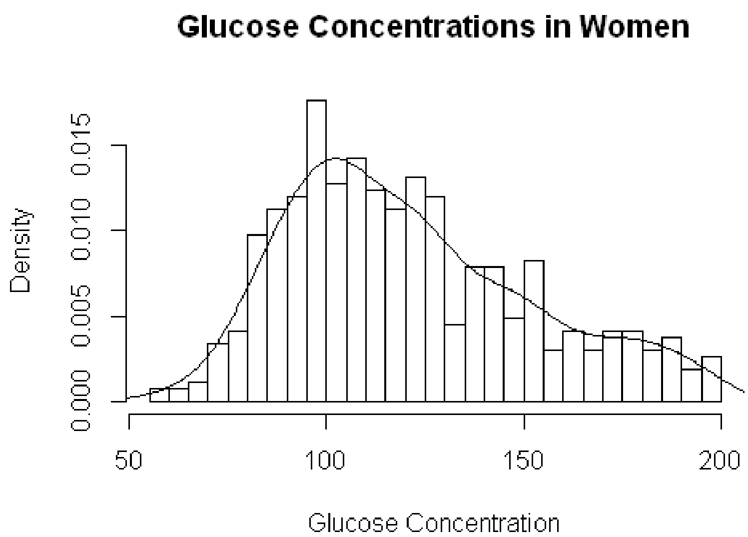

Figure 1: Problem 6.2a.

Without loss of generality Let X_i take values 0 or 1 with probability p and $1-p$ respectively. Therefore we have that

$$\begin{aligned}
P(X, p, \theta_1, \theta_2, \sigma_1^2, \sigma_2^2 | Y) &\propto P(Y|X, p, \theta_1, \theta_2, \sigma_1^2, \sigma_2^2) P(X, p, \theta_1, \theta_2, \sigma_1^2, \sigma_2^2) \\
&= \left(\frac{1}{\sqrt{2\pi\sigma_2^2}}\right)^{\sum x_i} \left(\frac{1}{\sqrt{2\pi\sigma_1^2}}\right)^{n-\sum x_i} \times \\
&\exp\{-\frac{1}{2\sigma_1^2} \sum_{x_i=0} (y_i - \theta_1)^2 - \frac{1}{2\sigma_2^2} \sum_{x_i=1} (y_i - \theta_2)^2\} \times \\
&P(X|p)P(p)P(\theta_1)P(\theta_2)P(\sigma_1^2)P(\sigma_2^2) \\
&= \left(\frac{1}{\sqrt{2\pi\sigma_2^2}}\right)^{\sum x_i} \left(\frac{1}{\sqrt{2\pi\sigma_1^2}}\right)^{n-\sum x_i} \times \\
&\exp\{-\frac{1}{2\sigma_1^2} \sum (1-x_i)(y_i - \theta_1)^2 - \frac{1}{2\sigma_2^2} \sum x_i (y_i - \theta_2)^2\} \times \\
&p^{n-\sum x_i}(1-p)^{\sum x_i} P(p)P(\theta_1)P(\theta_2)P(\sigma_1^2)P(\sigma_2^2)
\end{aligned}$$

Knowing that $P(X, p, \theta_1, \theta_2, \sigma_1^2, \sigma_2^2|Y) \propto P(X|p, \theta_1, \theta_2, \sigma_1^2, \sigma_2^2, Y)P(p, \theta_1, \theta_2, \sigma_1^2, \sigma_2^2|Y)$, we may isolate the terms with X in them. Hence we have

$$P(X|p, \theta_1, \theta_2, \sigma_1^2, \sigma_2^2, Y) \propto \left(\frac{1-p}{\sqrt{2\pi\sigma_2^2}}\right)^{\sum x_i} \left(\frac{p}{\sqrt{2\pi\sigma_1^2}}\right)^{n-\sum x_i} \exp\{\sum x_i \left(\frac{(y_i-\theta_1)^2}{2\sigma_1^2} - \frac{(y_i-\theta_2)^2}{2\sigma_2^2}\right)\}$$

We have that $P(X|p, \theta_1, \theta_2, \sigma_1^2, \sigma_2^2, Y) = \prod P(x_i|p, \theta_1, \theta_2, \sigma_1^2, \sigma_2^2, Y)$ by factorization. Thus we have

$$P(x_i|p,\theta_1,\theta_2,\sigma_1^2,\sigma_2^2,Y) \propto \left(\frac{1-p}{p}\frac{\sigma_1}{\sigma_2}\right)^{x_i} \exp\{x_i\left(\frac{(y_i-\theta_1)^2}{2\sigma_1^2}-\frac{(y_i-\theta_2)^2}{2\sigma_2^2}\right)\}$$
$$\propto \left(\frac{1-p}{p}\frac{\sigma_1}{\sigma_2}\exp\{\left(\frac{(y_i-\theta_1)^2}{2\sigma_1^2}-\frac{(y_i-\theta_2)^2}{2\sigma_2^2}\right)\}\right)^{x_i}$$

Thus the conditional x_i follows Bernoulli $\frac{1}{1+\frac{p-1}{p}\frac{\sigma_1}{\sigma_2}\exp\{-\left(\frac{(y_i-\theta_2)^2}{2\sigma_2^2}-\frac{(y_i-\theta_1)^2}{2\sigma_1^2}\right)\}}$.

Similarly we have $P(p|X,\theta_1,\theta_2,\sigma_1^2,\sigma_2^2,Y) \propto p^{n-\sum x_i}(1-p)^{\sum x_i}p^a(1-p)^b$. Thus the conditional is $beta(a+n-\sum x_i, b+\sum x_i)$.

Similarly we have

$$P(\theta_1|X,p,\theta_2,\sigma_1^2,\sigma_2^2,Y) \propto \exp\{-\frac{1}{2\sigma_1^2}\sum(1-x_i)(y_i-\theta_1)^2\}\frac{e^{-\frac{(\theta_1-\mu_0)^2}{2\tau_0^2}}}{\sqrt{2\pi\tau_0^2}}$$
$$\propto \exp\{-\frac{\theta_1^2(\sum(1-x_i)\tau_0^2+\sigma_1^2)-2\theta_1(\sum(1-x_i)y_i+\mu_0)}{2\sigma_1^2\tau_0^2}\}$$
$$\propto \exp\{-\frac{\sum(1-x_i)\tau_0^2+\sigma_1^2}{2\sigma_1^2\tau_0^2}(\theta_1-\frac{\sum(1-x_i)y_i+\mu_0}{\sum(1-x_i)\tau_0^2+\sigma_1^2})^2\}$$

Thus we have that the conditional distribution of θ_1 follows N$(\frac{\sum y_i(1-x_i)+\mu_0}{\sum(1-x_i)\tau_0^2+\sigma_1^2}, \frac{\sigma_1^2\tau_0^2}{\sum(1-x_i)\tau_0^2+\sigma_1^2})$.
Similarly we have

$$P(\theta_2|X,p,\theta_1,\sigma_1^2,\sigma_2^2,Y) \propto \exp\{-\frac{1}{2\sigma_2^2}\sum x_i(y_i-\theta_2)^2\}\frac{e^{-\frac{(\theta_2-\mu_0)^2}{2\tau_0^2}}}{\sqrt{2\pi\tau_0^2}}$$
$$\propto \exp\{-\frac{\theta_2^2(\sum x_i\tau_0^2+\sigma_2^2)-2\theta_2(\sum y_ix_i+\mu_0)}{2\sigma_2^2\tau_0^2}\}$$
$$\propto \exp\{-\frac{\sum x_i\tau_0^2+\sigma_2^2}{2\sigma_2^2\tau_0^2}(\theta_2-\frac{\sum y_ix_i+\mu_0}{\sum x_i\tau_0^2+\sigma_2^2})^2\}$$

Thus we have that the conditional distribution of θ_2 follows N$(\frac{\sum y_ix_i+\mu_0}{\sum x_i\tau_0^2+\sigma_2^2}, \frac{\sigma_2^2\tau_0^2}{\sum x_i\tau_0^2+\sigma_2^2})$.
Similarly we have

$$P(\sigma_1^2|X,p,\theta_1,\theta_2,\sigma_2^2,Y) \propto \left(\frac{1}{\sqrt{2\pi\sigma_1^2}}\right)^{n-\sum x_i}\exp\{-\frac{1}{2\sigma_1^2}\sum(1-x_i)(y_i-\theta_1)^2\}$$
$$\times (\sigma_1^2)^{-(\nu_0/2+1)}\exp\{-\nu_0\sigma_0^2/2/\sigma_1^2\}$$
$$\propto (\sigma_1^2)^{-(n-\sum x_i+\nu_0)/2-1}\exp\{-\frac{1}{2\sigma_1^2}(\sum(1-x_i)(y_i-\theta_1)^2+\nu_0\sigma_0^2)\}$$

which implies the conditional distribution is Inverse-Gamma$((n-\sum x_i+\nu_0)/2, (\sum(1-x_i)(y_i-\theta_1)^2+\nu_0\sigma_0^2)/2)$

Similarly we have

$$P(\sigma_2^2|X,p,\theta_1,\theta_2,\sigma_1^2,Y) \propto \left(\frac{1}{\sqrt{2\pi\sigma_2^2}}\right)^{\sum x_i} \exp\{-\frac{1}{2\sigma_2^2}\sum x_i(y_i-\theta_2)^2\}$$
$$\times (\sigma_2^2)^{-(\nu_0/2+1)}\exp\{-\nu_0\sigma_0^2/2/\sigma_2^2\}$$
$$\propto (\sigma_2^2)^{-(\sum x_i+\nu_0)/2-1}\exp\{-\frac{1}{2\sigma_2^2}(\sum x_i(y_i-\theta_2)^2+\nu_0\sigma_0^2)\}$$

which implies the conditional distribution is Inverse-Gamma$((\sum x_i+\nu_0)/2, (\sum x_i(y_i-\theta_2)^2+\nu_0\sigma_0^2)/2)$.

c) Setting $a=b=1, \mu_0=120, \tau_0^2=200, \sigma_0^2=1000$ and $\nu_0=10$, implement the Gibbs sampler for at least 10,000 iterations. Let $\theta_{(1)}^{(s)}=\min\{\theta_1^{(s)},\theta_2^{(s)}\}$ and $\theta_{(2)}^{(s)}=\max\{\theta_1^{(s)},\theta_2^{(s)}\}$. Compute and plot the autocorrelation functions of $\theta_{(1)}^{(s)}$ and $\theta_{(2)}^{(s)}$, as well as their effective sample sizes.

I ran into an issue where the latent probability flips between 0 and 1 after the first iteration. Thus everything is out of alignment. I suspect an error because I used the variance instead of precision. We see in Figures 2 and 3 that we have almost no correlation after the first step.

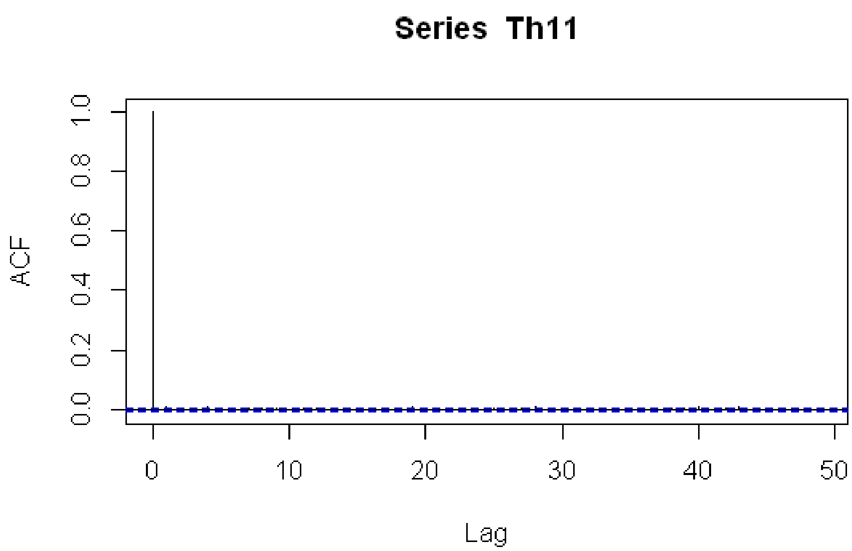

Figure 2: Autocorrelation for $\theta_{(1)}$

d) For each iteration s of the Gibbs sampler, sample a value $x \sim binary(p^{(s)})$, then sample $\bar{Y}^{(s)} \sim normal(\theta_x^{(s)}, \sigma_x^{2(s)})$. Plot a histogram or kernel density estimate for the empirical distribution of $\bar{Y}^{(1)},\ldots,\bar{Y}^{(S)}$, and compare the distribution in part a). Discuss the adequacy of this two-component mixture model for the glucose data.

In Figure 4, we see that the mixing did not get close the the original data set. Again I believe it is a coding issue.

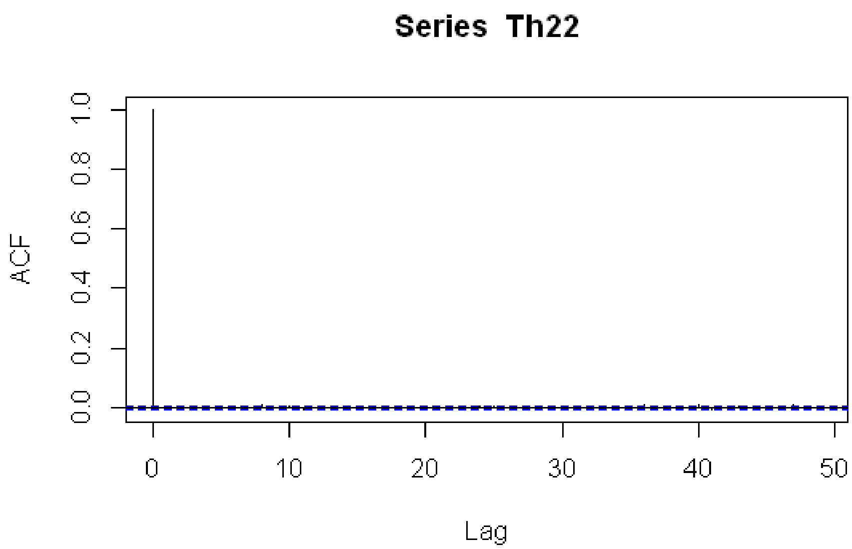

Figure 3: Autocorrelation for $\theta_{(2)}$

```
#PRoblem 1
getwd()
setwd('m9810Fall2014')
school1=scan("school1.dat")
school2=scan("school2.dat")
school3=scan("school3.dat")
#use books definitions
k_0=1
mu_0=5
s2_0=4
nu_0=2
n.1=length(school1)
n.2=length(school2)
n.3=length(school3)
#posterior variables
k_n.1=k_0+n.1
k_n.2=k_0+n.2
k_n.3=k_0+n.3
nu_n.1=nu_0+n.1
nu_n.2=nu_0+n.2
nu_n.3=nu_0+n.3
ybar.1=mean(school1)
s2.1=var(school1)
ybar.2=mean(school2)
s2.2=var(school2)
ybar.3=mean(school3)
s2.3=var(school3)
s2_n.1 = (nu_0*s2_0 + (n.1-1)*s2.1 + k_0*n.1*(ybar.1-mu_0)^2/(k_n.1))/nu_n.1
mu_n.1=(n.1*ybar.1+k_0*mu_0)/k_n.1
```

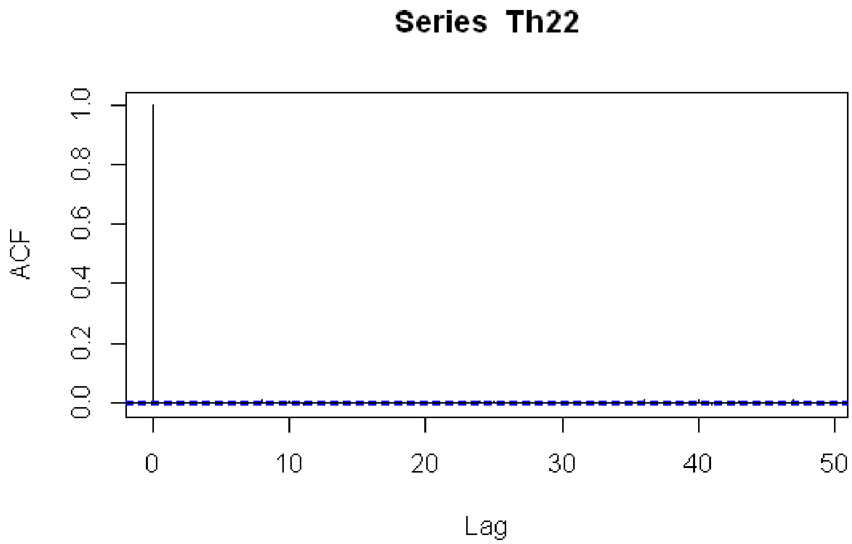

Figure 4: Posterior predictive for $\bar{Y}$

```
s2_n.2 = (nu_0*s2_0 + (n.2-1)*s2.2 + k_0*n.1*(ybar.2-mu_0)^2/(k_n.2))/nu_n.2
mu_n.2=(n.2*ybar.2+k_0*mu_0)/k_n.2
s2_n.3 = (nu_0*s2_0 + (n.3-1)*s2.3 + k_0*n.3*(ybar.3-mu_0)^2/(k_n.3))/nu_n.3
mu_n.3=(n.3*ybar.3+k_0*mu_0)/k_n.3
#monte carlo
iter=1e6
postsigma2.1 = 1/rgamma(iter,nu_n.1/2,rate=s2_n.1*nu_n.1/2)
posttheta.1 = rnorm(iter,mu_n.1,(postsigma2.1/k_n.1)^.5)
newy.1=rnorm(iter,posttheta.1,(postsigma2.1/k_n.1)^.5)
postsigma2.2 = 1/rgamma(iter,nu_n.2/2,rate=s2_n.2*nu_n.2/2)
posttheta.2 = rnorm(iter,mu_n.2,(postsigma2.2/k_n.2)^.5)
newy.2=rnorm(iter,posttheta.2,(postsigma2.2/k_n.2)^.5)
postsigma2.3 = 1/rgamma(iter,nu_n.3/2,rate=s2_n.3*nu_n.3/2)
posttheta.3 = rnorm(iter,mu_n.3,(postsigma2.3/k_n.3)^.5)
newy.3=rnorm(iter,posttheta.3,(postsigma2.3/k_n.3)^.5)

#mean of posterior theta and 95%
mean(posttheta.1)
mean(posttheta.2)
mean(posttheta.3)
quantile(posttheta.1,c(.025,.0975))
quantile(posttheta.2,c(.025,.0975))
quantile(posttheta.3,c(.025,.0975))

#mean of sigma
mean((postsigma2.1)^.5)
mean((postsigma2.2)^.5)
mean((postsigma2.3)^.5)
quantile((postsigma2.1)^.5,c(.025,.0975))
```

```r
quantile((postsigma2.2)^.5,c(.025,.0975))
quantile((postsigma2.3)^.5,c(.025,.0975))

#P(thi<thj<thk)
pth=rep(NA,6)
#123
pth[1]=sum(posttheta.1<posttheta.2 & posttheta.2 < posttheta.3)/iter
#132
pth[2]=sum(posttheta.1<posttheta.3 & posttheta.3 < posttheta.2)/iter
#213
pth[3]=sum(posttheta.2<posttheta.1 & posttheta.1 < posttheta.3)/iter
#231
pth[4]=sum(posttheta.2<posttheta.3 & posttheta.3 < posttheta.1)/iter
#312
pth[5]=sum(posttheta.3<posttheta.1 & posttheta.1 < posttheta.2)/iter
#321
pth[6]=sum(posttheta.3<posttheta.2 & posttheta.2 < posttheta.1)/iter
pth

#P(yi<yj<yk)
pnewy=rep(NA,6)
#123
pnewy[1]=sum(newy.1<newy.2 & newy.2 < newy.3)/iter
#132
pnewy[2]=sum(newy.1<newy.3 & newy.3 < newy.2)/iter
#213
pnewy[3]=sum(newy.2<newy.1 & newy.1 < newy.3)/iter
#231
pnewy[4]=sum(newy.2<newy.3 & newy.3 < newy.1)/iter
#312
pnewy[5]=sum(newy.3 <newy.1 & newy.1 < newy.2)/iter
#321
pnewy[6]=sum(newy.3 <newy.2 & newy.2 < newy.1)/iter
pnewy

#y1 biggest
pnewy[6]+pnewy[4]
#th1 biggest
pth[6]+pth[4]

###Problem no 2
setA=scan('menchild30bach.dat')
setB=scan('menchild30nobach.dat')
yA=sum(setA)
yB=sum(setB)
nA=length(setA)
nB=length(setB)
ath=2
bth=1
abg=c(8,16,32,64,128)
```

```
## Use Gibbs sampler to sample theta and gamma
iter = 5e4; ## number of iterations (i.e., T)

## Save records
Th = rep(NA, iter);
Ga = rep(NA, iter);
ThM= rep(NA,5);
GaM= rep(NA,5);
FinalM=rep(NA,5);
for(i in 1:5){
#initiate sampler
th=rgamma(1,2,rate=1)
ga=rgamma(1,abg[i],rate=abg[i])

Th[1]=th
Ga[1]=ga
## Start the Gibbs sampler
for(t in 2:iter){
  #use previous generations ga
  th = rgamma(1,yA+yB+ath  ,rate=bth+nA+nB*ga);
  #use this generations th
  ga = rgamma(1, yB+abg[i]  ,rate=abg[i]+nB*th);

  Th[t] = th;
  Ga[t] = ga;
}

#plot(Th, typ = 'l')
#plot(Ga, typ = 'l')

ThM[i]=mean(Th[-(1:500)])
GaM[i]=mean(Ga[-(1:500)])
FinalM[i]=sum(Th[-(1:500)]*(Ga[-(1:500)]-1))/(iter-500)
}
FinalM
plot(1:5,FinalM)

###Problem no 3
glu=scan('glucose.dat')
hist(glu, prob = TRUE,xlab="Glucose Concentration", main = "Glucose Concentrations in
dglu<-density(glu)
lines(dglu)
#initial variables
a=1
b=1
mu_0=120
t2_0=200
s2_0=1000
nu_0=10
```

```
n=length(glu)
ybar=mean(glu)
s2=var(glu)
#number of samples
iter=1e5
#storage
Th.1=rep(NA,iter)
Th.2=rep(NA,iter)
S2.1=rep(NA,iter)
S2.2=rep(NA,iter)
P=rep(NA,iter)
X=matrix(,nrow=n,ncol=iter)
newy=rep(NA,iter)
#initiate sampler
p=rbeta(1,a,b)
P[1]=p
newx=rbinom(1,1,p)
x=rbinom(n,1,p)
#x=sample(c(0,1),replace=TRUE,n,prob=c(p[1],1-p[1]))
X[,1]=x
s2.1=1/rgamma(1,nu_0/2,rate=nu_0*s2_0/2)
S2.1[1]=s2.1
s2.2=1/rgamma(1,nu_0/2,rate=nu_0*s2_0/2)
S2.2[1]=s2.2
th.1=rnorm(1,mu_0,t2_0)
th.2=rnorm(1,mu_0,t2_0)
Th.1[1]=min(th.1,th.2)
Th.2[1]=max(th.1,th.2)
if (newx==1){
  newy[1]=rnorm(1,th.1,sd=sqrt(s2.1))
}else
{
  newy[1]=rnorm(1,th.2,sd=sqrt(s2.2))
}
for(i in 2:3){
  p<-rbeta(1,a+n-sum(x),b+sum(x))
  for(j in 1:n){
    pp1<-p*exp(-((glu[j]-th.1)^2)/(2*s2.1))/sqrt(s2.1);
    pp2<-(1-p)*exp(-((glu[j]-th.2)^2)/(2*s2.2))/sqrt(s2.2);
    ptemp=pp1/(pp1+pp2);
    print(ptemp)
    x[j]=rbinom(1,1,ptemp)
  }
  s2.1=1/rgamma(1,(n-sum(x)+nu_0)/2,rate=(sum((1-x)*(glu-th.1)^2)+nu_0*s2_0)/2)
  s2.2=1/rgamma(1,(sum(x)+nu_0)/2,rate=(sum(x*(glu-th.2)^2)+nu_0*s2_0)/2)
  th.1=rnorm(1,(sum(glu*(1-x))+mu_0)/(sum(1-x)*t2_0+s2.1),sd=sqrt(s2.1*t2_0/(sum(1-x)
  th.2=rnorm(1,(sum(glu*x)+mu_0)/(sum(x)*t2_0+s2.2),sd=sqrt(s2.2*t2_0/(sum(x)*t2_0+s2


  P[i]=p
```

```
  X[,i]=x
  S2.1[i]=s2.1
  S2.2[i]=s2.2
  Th.1[i]=min(th.1,th.2)
  Th.2[i]=max(th.1,th.2)
}
for (i in 2:iter){
  newx<-rbinom(1,1,P[i])
  if (newx==0){
    newy[i]=rnorm(1,Th.1[i],sd=sqrt(S2.1[i]))
  }else
  {
    newy[i]=rnorm(1,Th.2[i],sd=sqrt(S2.2[i]))
  }
}

library('coda')
Th11=as.mcmc(Th.1[-(1:1000)])
traceplot(Th11)
acf(Th11)
Th22=as.mcmc(Th.2[-(1:1000)])
traceplot(Th22)
acf(Th22)
hist(newy, prob = TRUE, xlab="Glucose Concentration", main = "Glucose Concentrations i
dnewy<-density(newy)
lines(dnewy)
NEWYY=as.mcmc(newy)
traceplot(NEWYY)
```