Final Answers

2018150420 정해원

2022-12-22

Question 1

(a)

$$\begin{aligned} Y_i|\theta_1, \sigma^2 &\sim N(\theta_1, \sigma^2), \;\; i = 1, \cdots, m \\ Y_i|\theta_1, \theta_2, \sigma^2 &\sim N(\theta_1 + \theta_2, \sigma^2), \;\; i = m+1, \cdots, n \\ \theta_1 &\sim N(0, 100^2) \\ \theta_2 &\sim N(0, 100^2) \\ \sigma^2 &\sim IG(0.01, 0.01) \end{aligned}$$

Then the joint posterior distribution can be calculated by

$$\begin{aligned} p(\theta_1, \theta_2, \sigma^2|Y_i) \propto & \prod_{i=1}^{m} \frac{1}{\sigma} exp(-\frac{1}{2\sigma^2}(y_i - \theta_1)^2) \prod_{i=m+1}^{n} \frac{1}{\sigma} exp(-\frac{1}{2\sigma^2}(y_i - \theta_1 - \theta_2)^2) \times \\ & \prod_{j=1}^{2} \frac{1}{\sigma} exp(-\frac{1}{2\sigma^2}(y_i - \theta_1)^2) \prod_{j=1}^{2} \frac{1}{\sigma} exp(-\frac{1}{2\sigma^2}(\theta_j)^2) \times (\sigma^2)^{-\frac{1}{100}-1} e^{-\frac{1}{100\sigma^2}} \end{aligned}$$

The full conditional distribution of θ_1 is

$$p(\theta_1|\theta_2, \sigma^2, Y_i6) \propto \prod_{i=1}^{m} \frac{1}{\sigma} exp(-\frac{1}{2\sigma^2}(y_i - \theta_1)^2) \prod_{i=m+1}^{n} \frac{1}{\sigma} exp(-\frac{1}{2\sigma^2}(y_i - \theta_1 - \theta_2)^2) \times exp(-\frac{\theta_1^2}{2\sigma^2})$$

Then, $\frac{1}{\tau_1^2} = \frac{1}{\sigma^2} + \frac{m}{\sigma^2} + \frac{n-m}{\sigma^2}$ so $\tau_1^2 = \frac{\sigma^2}{n+1}$ and $\mu_1 = \frac{\sigma^2}{n+1}(\frac{0}{\sigma^2} + \frac{m\bar{y_1}}{\sigma^2} + \frac{(n-m)\bar{y_2}}{\sigma^2})$ This leads to

$$\theta_1|\theta_2, \sigma^2, Y_i \sim N(\frac{m\bar{y_1} + (n-m)\bar{y_2}}{n+1}, \frac{\sigma^2}{n+1})$$

The full conditional distribution of θ_2 is

$$p(\theta_2|\theta_1, \sigma^2, Y_i) \propto \prod_{i=m+1}^{n} \frac{1}{\sigma} exp(-\frac{1}{2\sigma^2}(y_i - \theta_1 - \theta_2)^2) \times exp(-\frac{\theta_2^2}{2\sigma^2})$$

By similar calculation, $\tau_2^2 = \frac{\sigma^2}{n-m+1}$, $mu_2 = \frac{(n-m)\bar{y_2}}{n-m+1}$ and this leads to

$$\theta_2|\theta_1, \sigma^2, Y_i \sim N(\frac{(n-m)\bar{y_2}}{n-m+1}, \frac{\sigma^2}{n-m+1})$$

The full conditional distribution of σ^2 uses all information from hyperprior, prior, likelihood.

$$\begin{aligned} p(\sigma^2|\theta_1, \theta_2, Y_i) \propto & \prod_{i=1}^{m} \frac{1}{\sigma} exp(-\frac{1}{2\sigma^2}(y_i - \theta_1)^2) \prod_{i=m+1}^{n} \frac{1}{\sigma} exp(-\frac{1}{2\sigma^2}(y_i - \theta_1 - \theta_2)^2) \times \\ & \prod_{j=1}^{2} \frac{1}{\sigma} exp(-\frac{1}{2\sigma^2}(y_i - \theta_1)^2) \prod_{i=m+1}^{n} \frac{1}{\sigma} exp(-\frac{1}{2\sigma^2}(\theta_j)^2) \times (\sigma^2)^{-\frac{1}{100}-1} e^{-\frac{1}{100\sigma^2}} \\ \propto & (\sigma^2)^{-\frac{n}{2}-\frac{1}{100}-2} exp(-\frac{1}{2\sigma^2}(\sum_{i=1}^{m}(y_i - \theta_1)^2 + \sum_{i=m+1}^{n}(y_i - \theta_1 - \theta_2)^2 + \frac{\theta_1^2 + \theta_2^2}{2})) \end{aligned}$$

this leads to

$$\sigma^2|\theta_1, \theta_2, Y_i \sim IG(0.01 + \frac{n+2}{2}, \; 0.01 + \frac{1}{2\sigma^2}(\sum_{i=1}^{m}(y_i - \theta_1)^2 + \sum_{i=m+1}^{n}(y_i - \theta_1 - \theta_2)^2 + \frac{\theta_1^2 + \theta_2^2}{2}))$$

(b)

```
library(extraDistr)
m<-50; n<-100; theta1<-10; theta2<-10; sigma<-2; sigma2<-(sigma)^2
x<-c(rnorm(m, theta1, sigma), rnorm(n-m, theta2, sigma)) # data
mean.x1<-mean(x[1:m]); mean.x2<-mean(x[(n-m+1):n]); var.x<-var(x)
S<-10^4
PHI<-matrix(nrow=S,ncol=3)
sigma2<-(sigma)^2
PHI[1,1] <- mean.x1
PHI[1,2] <- mean.x2
PHI[1,3] <- var.x
### Gibbs sampling
for(s in 2:S) {
  PHI[s,1]<- rnorm(1, (m*PHI[s-1,1]+(n-m)*PHI[s-1,2])/(n+1), sqrt(PHI[s-1,3]/(n+1)))
  PHI[s,2]<- rnorm(1, ((n-m)*PHI[s-1,2]/(n+1)), sqrt(PHI[s-1,3]/(n-m+1)))
  PHI[s,3] <- rinvgamma(1, 0.01+(n+2)/2, 0.01+0.5*(m*(-PHI[s-1,1])^2+(n-m)*(0-PHI[s-1,1]-PHI[s-1,2])^2+(PHI[s-1,1]^2+
PHI[s-1,2]^2)/2))
}
PHI[,3] <- sqrt(PHI[,3])
head(PHI,20)
```

```
##              [,1]         [,2]        [,3]
##  [1,] 10.151453885 10.223987872  2.05662094
##  [2,] 10.118278207  5.322156482 16.66599181
##  [3,]  6.662260234  5.593676588 12.36338771
##  [4,]  4.313587022  1.769361542  9.93592211
##  [5,]  4.595375968  1.255995425  4.76876647
##  [6,]  3.115435383  0.450017974  5.73247759
##  [7,]  1.954418744  0.027834314  3.55020949
##  [8,]  1.376619720 -0.517312267  1.73569147
##  [9,]  0.320569057  0.021009095  1.26881405
## [10,]  0.279296792  0.212185872  0.33319066
## [11,]  0.217391961  0.046365417  0.43409005
## [12,]  0.086738394 -0.020882881  0.22580063
## [13,]  0.035598500 -0.013852208  0.07981754
## [14,]  0.018735597 -0.024276946  0.03782358
## [15,]  0.003213718 -0.023570672  0.01981026
## [16,] -0.011201669 -0.015403163  0.01906899
## [17,] -0.017985015 -0.010160218  0.02694355
## [18,] -0.017098450 -0.003505270  0.02700269
## [19,] -0.014309634 -0.001221099  0.02099669
## [20,] -0.006453049  0.001846775  0.02174384
```

The means of $\theta_1, \theta_2, \sigma$ are in below in order.

```
means <- apply(PHI[,], 2, mean)
means
```

```
## [1] 0.004367541 0.002425107 0.020339211
```

The medians of $\theta_1, \theta_2, \sigma$ are in below in order.

```
apply(PHI[,], 2, median)
```

```
## [1]  3.908921e-05 -1.010840e-05  1.436219e-02
```

The 95% CIs of $\theta_1, \theta_2, \sigma$ are shown as a matrix form below in order.

```
sds<-apply(PHI[,], 2, sd); CI<-matrix(nrow=3, ncol=2)
for (i in 1:3) {CI[i,] <- c(quantile(PHI[,i], 0.025), quantile(PHI[,i], 0.975))}
rownames(CI) <- c("theta1", "theta2", "sigma"); CI
```

```
##                [,1]        [,2]
## theta1 -0.004531196 0.004759987
## theta2 -0.004521290 0.004666800
## sigma   0.012475430 0.016908707
```

(c)

Using

$$\begin{aligned} Y_i|\theta_1, \sigma^2 &\sim N(\theta_1, \sigma^2), \;\; i = 1, \cdots, m \\ Y_i|\theta_1, \theta_2, \sigma^2 &\sim N(\theta_1 + \theta_2, \sigma^2), \;\; i = m+1, \cdots, n \\ \theta_1 &\sim N(0, 100^2) \\ \theta_2 &\sim N(0, 100^2) \\ \sigma^2 &\sim IG(0.01, 0.01) \end{aligned}$$

and another condition,

$$m \sim Unif(1, \cdots, n)$$

then

$$\begin{aligned} p(m|\theta_1, \theta_2, \sigma^2, Y_i) &\propto \prod_{i=1}^{m} \frac{1}{\sigma} exp(-\frac{1}{2\sigma^2}(y_i - \theta_1)^2) \prod_{i=m+1}^{n} \frac{1}{\sigma} exp(-\frac{1}{2\sigma^2}(y_i - \theta_1 - \theta_2)^2)) \\ &\propto \frac{m}{100}(\frac{1}{\sigma} exp(-\frac{1}{2\sigma^2}(y_i - \theta_1)^2))^m \frac{n-m}{100}(\frac{1}{\sigma} exp(-\frac{1}{2\sigma^2}(y_i - \theta_1 - \theta_2)^2))^{n-m} \end{aligned}$$

which leads to

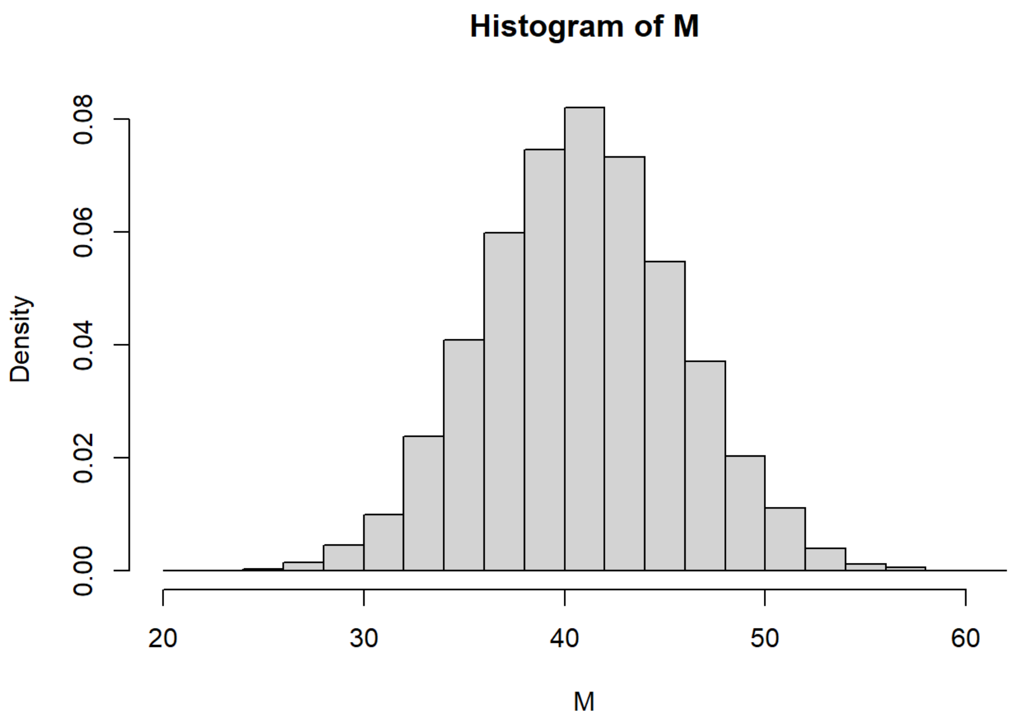
$$m|\theta_1, \theta_2, \sigma^2, Y_i \sim Binom(n, \frac{m(\frac{1}{\sigma} exp(-\frac{1}{2\sigma^2}(y_i - \theta_1)^2))^m}{m(\frac{1}{\sigma} exp(-\frac{1}{2\sigma^2}(y_i - \theta_1)^2))^m + (n-m)(\frac{1}{\sigma} exp(-\frac{1}{2\sigma^2}(y_i - \theta_1 - \theta_2)^2))^{n-m}})$$

```
m.post <- m*prod(x[1:m])/(m*prod(x[1:m])+(n-m)*prod(x[(n-m+1):n]))
M <- rbinom(S, n, m.post)
mean(M)
```

```
## [1] 41.3629
```

The histogram is shown as below.

```
hist(M, freq=F)
```



Question 2

(a)

Since $p(X) \propto \theta^x(1-\theta)^{n-x}$, $\theta|\eta \propto \theta^{\eta-1}$, the conditional posterior is $p(\theta|\eta, x) \propto \theta^{\eta-1+x}(1-\theta)^{n-x}$ and this means that

$$\theta_n|\eta_{n-1}, x \sim Beta(\eta_{n-1} + x, n - x + 1)$$

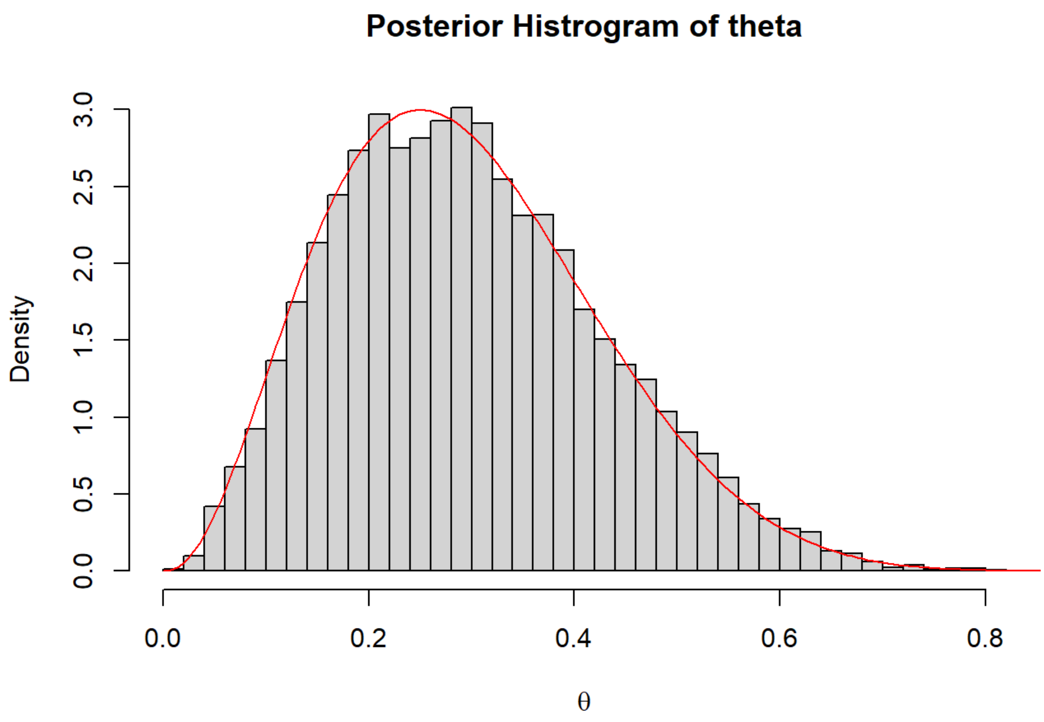
Also, $p(\eta|x, \theta) \propto \theta^{\eta-1}e^{-a\eta}$, and this means that

$$\eta_n|x, \theta_n \sim \Gamma(\theta_n + 1, a)$$

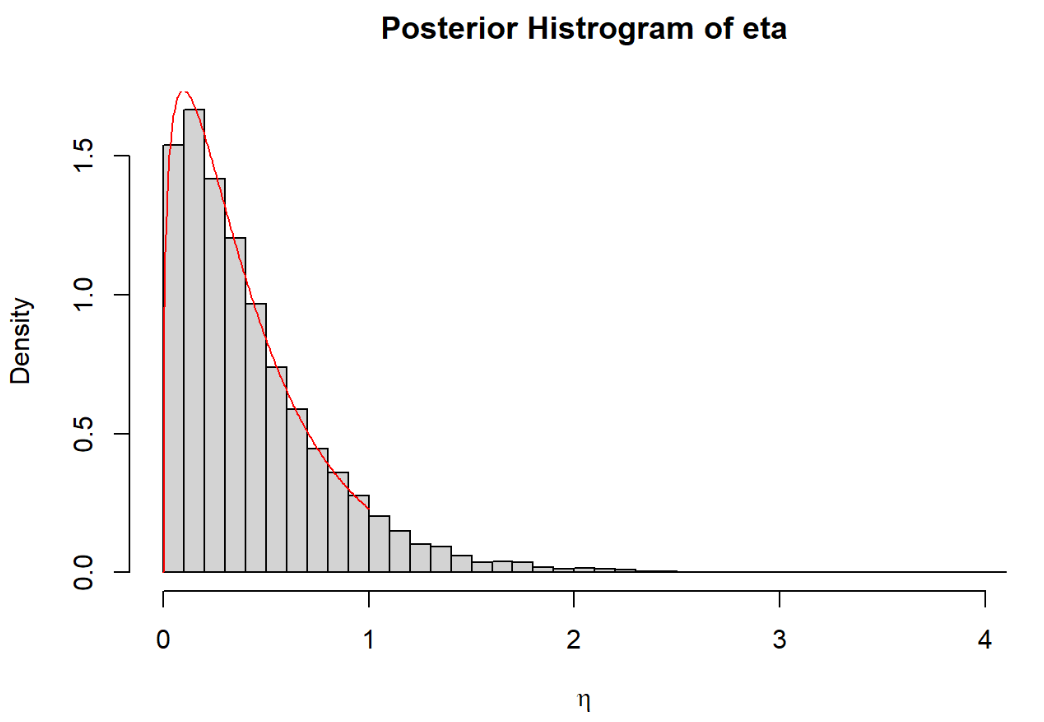
Simulating by $x = 3, n = 10, a = 3$

```
n <- 10; x<-3; a<-3
S<-10^4
PHI<-matrix(nrow=S,ncol=2)
b <- rgamma(1, shape=1, rate=a)
PHI[1,1] <- rbeta(1, b+x, n-x+1)
PHI[1,2] <- rgamma(1, shape=1, rate=a+PHI[1,1])
### Gibbs sampling
for(s in 2:S) {
  PHI[s,1]<- rbeta(1, PHI[s-1,2]+x, n-x+1)
  PHI[s,2]<- rgamma(1, shape=1+PHI[s,1], rate=a)
}

grid<-seq(0, 1, by=0.005)
hist(PHI[,1], breaks=40, freq=F, main=paste("Posterior Histrogram of", expression(theta)), xlab=expression(theta))
lines(grid, dbeta(grid, 10/3, 8), col="red")
```




```
hist(PHI[,2], breaks=40, freq=F, main=paste("Posterior Histrogram of", expression(eta)), xlab=expression(eta))
lines(grid, dgamma(grid, shape=1+5/17, rate=3), col="red")
```



(b)

Again, we do the Gibbs sampling.

```
n <- 10; x<-3; a<-3
S<-10^4
PHI<-matrix(nrow=S,ncol=2)
b <- rgamma(1, shape=1, rate=a)
PHI[1,1] <- rbeta(1, b+x, n-x+1)
PHI[1,2] <- rgamma(1, shape=1, rate=a+PHI[1,1])
### Gibbs sampling
for(s in 2:S) {
  PHI[s,1]<- rbeta(1, PHI[s-1,2]+x, n-x+1)
  PHI[s,2]<- rgamma(1, shape=1+PHI[s,1], rate=a)
}
```

Now, we will make a squared error function for θ, η

```
posterior_risk_theta <- function(a,theta){
  losses_a <- (theta-a)^2
  return(mean(losses_a))
}
posterior_risk_eta <- function(b,eta){
  losses_b <- (eta-b)^2
  return(mean(losses_b))
}
```

The Bayes estimator by squared error function can be estimated as following

```
a <- seq(0,1,by=0.001)
post_risk_theta <- apply(as.matrix(a),1,function(a) posterior_risk_theta(a, PHI[,1]))
b <- seq(0,5, by=0.01)
post_risk_eta <- apply(as.matrix(b),1,function(b) posterior_risk_eta(b,PHI[,2]))

B_theta <- a[which.min(post_risk_theta)];B_eta <- a[which.min(post_risk_eta)]
print(round(B_theta,4)); print(round(B_eta, 4))
```

```
## [1] 0.302
```

```
## [1] 0.044
```

It is quite similar with the sample mean of each parameters from Gibbs sampling

```
paste("Sample mean of",expression(theta),":",round(mean(PHI[,1]),5))
```

```
## [1] "Sample mean of theta : 0.30153"
```

```
paste("Sample mean of",expression(eta),":",round(mean(PHI[,2]),5))
```

```
## [1] "Sample mean of eta : 0.43921"
```

Question 3

(a)

Given the distributions of parameters and data, those can be expressed as

$$
\begin{gathered}
Y_i|x_i = 1, \lambda, \gamma, \beta \sim Poisson(\lambda) \\
Y_i|x_i = 2, \lambda, \gamma, \beta \sim Poisson(\gamma\lambda) \\
\lambda \sim \Gamma(0.1, \beta),\ \gamma \sim \Gamma(0.1, \beta),\ \beta \sim \Gamma(0.1, 1)
\end{gathered}
$$

Then the joint posterior distribution can be calculated by

$$
p(\lambda, \gamma, \beta, p, x_i|Y_i) \propto \prod_{i=1}^{n}(\frac{e^{-\lambda}\lambda^{y_i}}{y_i!})^{1-x_i} \prod_{i=1}^{n}(\frac{e^{-\gamma\lambda}(\gamma\lambda)^{y_i}}{y_i!})^{x_i} \times \beta^{0.1}\lambda^{-0.9}exp(-\beta\lambda) \times \beta^{0.1}\gamma^{-0.9}exp(-\beta\gamma) \times \beta^{-0.9}exp(-\beta)
$$

The full conditional distribution of λ is

$$p(\lambda|y_1,\cdots,y_n,x_1,\cdots,x_n,\gamma,\beta) \propto \prod_{i=1}^{n}(\frac{e^{-\lambda}\lambda^{y_i}}{y_i!})^{1-x_i}\prod_{i=1}^{n}(\frac{e^{-\gamma\lambda}(\gamma\lambda)^{y_i}}{y_i!})^{x_i} \times \beta^{0.1}\lambda^{-0.9}exp(-\beta\lambda)$$
$$\propto \lambda^{\sum y_i - 0.9}exp(\lambda(n-\sum x_i) - \gamma\lambda\sum x_i - \beta\lambda))$$
$$\lambda|y_1,\cdots,y_n,x_1,\cdots,x_n,\gamma,\beta \sim \Gamma(0.1+\sum y_i, \beta+n+(\gamma-1)\sum x_i)$$

The full conditional distribution of γ is

$$p(\gamma|y_1,\cdots,y_n,x_1,\cdots,x_n,\lambda,\beta) \propto \prod_{i=1}^{n}(\frac{e^{-\gamma\lambda}(\gamma\lambda)^{y_i}}{y_i!})^{x_i} \times \gamma^{-0.9}exp(-\beta\gamma) \propto \frac{\gamma^{\sum y_i(1-x_i)-0.9}exp(-(\beta+\lambda(n-\sum x_i))\gamma)}{\prod y_i!}$$
$$\gamma|y_1,\cdots,y_n,x_1,\cdots,x_n,\lambda,\beta \sim \Gamma(1+\sum y_i(1-x_i), \beta+\lambda(n-\sum x_i))$$

The full conditional distribution of β is

$$p(\beta|y_1,\cdots,y_n,x_1,\cdots,x_n,\lambda,\gamma) \propto \beta^{0.1}\lambda^{-0.9}exp(-\beta\lambda) \times \beta^{0.1}\gamma^{-0.9}exp(-\beta\gamma) \times \beta^{-0.9}exp(-\beta) \propto \beta^{-0.7}exp(-(1+\lambda+\gamma)\beta)$$
$$p(\beta|y_1,\cdots,y_n,x_1,\cdots,x_n,\lambda,\gamma) \sim \Gamma(0.3, 1+\lambda+\gamma)$$

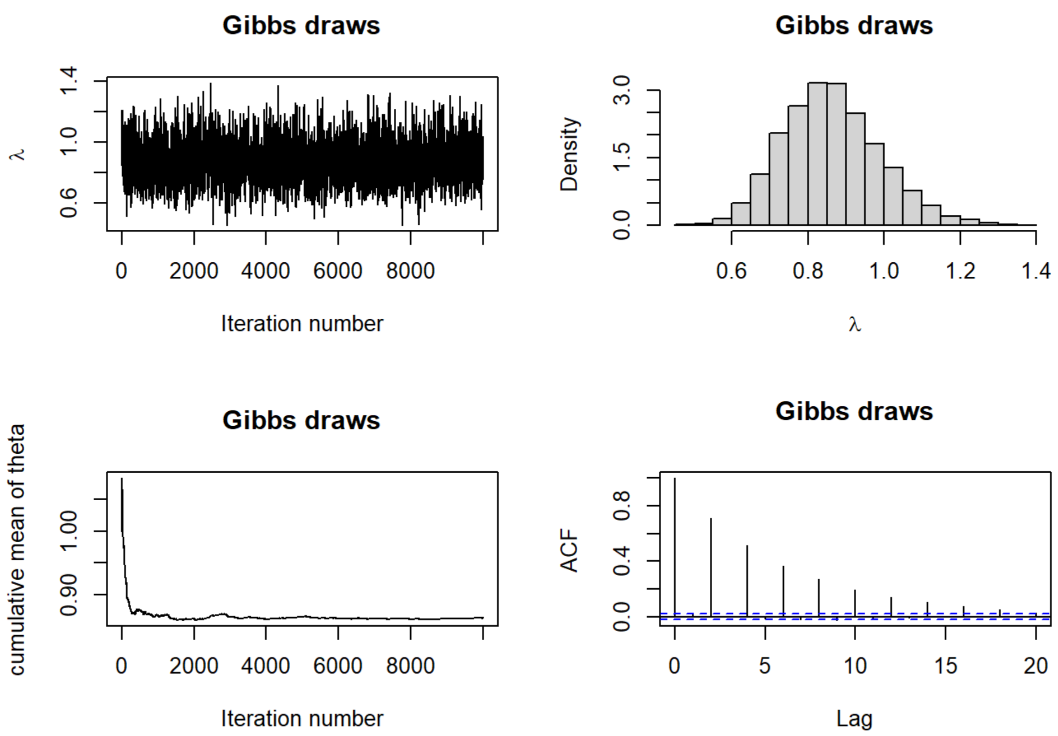
(b)

```
n<-100; m<-50; L<-1; G<-2
y <- c(rpois(m, lambda=L), rpois(n-m, lambda=L*G));
y1 <- y[1:m]; y2 <- y[(n-m+1):n]
PHI<-matrix(nrow=S,ncol=3)
PHI[1,1] <- L
PHI[1,2] <- G
PHI[1,3] <- rgamma(1, shape=0.1, rate=1)
s<-2
### Gibbs sampling
for(s in 2:S) {
  PHI[s,1]<- rgamma(1,shape=0.1+sum(y), rate=PHI[s-1,3]+n+(PHI[s-1,2]-1)*m)
  PHI[s,2]<- rgamma(1,shape=0.1+sum(y2), rate=PHI[s-1,3]+PHI[s-1,1]*(n-m))
  PHI[s,3] <- rgamma(1, shape=0.3, rate=1+PHI[s-1,1]+PHI[s-1,2])
}
```

Now the Gibbs samples of λ, γ, β are made. To see whether they are mixed well, we will check several graphs including ACF. First, for the λ,

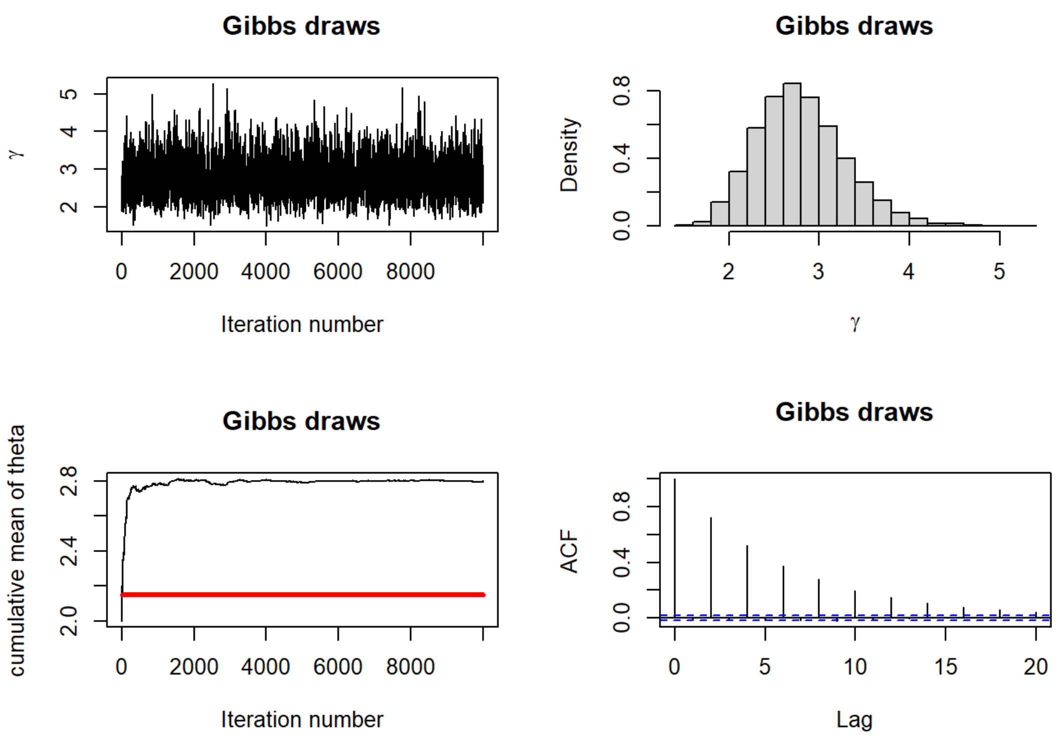
```
library(mnormt)
par(mfrow=c(2,2))

plot(PHI[,1], type="l", main='Gibbs draws',
     xlab="Iteration number", ylab=expression(lambda))
hist(PHI[,1], freq = FALSE, main='Gibbs draws', breaks=20,xlab=expression(lambda))
plot(cumsum(PHI[,1])/seq(1,S),type="l",main='Gibbs draws',
     xlab='Iteration number', ylab='cumulative mean of theta')
lines(seq(1,S),1*matrix(1.1,1,S),col="red",lwd=3)
acf(PHI[,1], main='Gibbs draws', lag.max = 20)
```



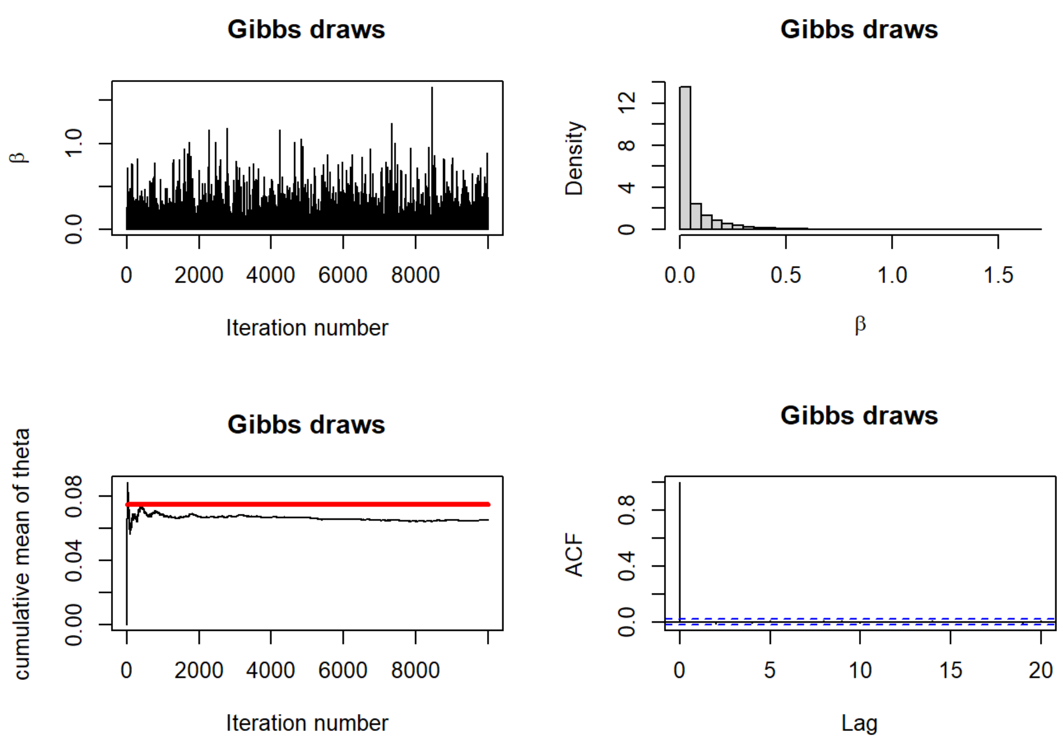
For, γ

```
par(mfrow=c(2,2))
plot(PHI[,2], type="l", main='Gibbs draws',
     xlab="Iteration number", ylab=expression(gamma))
hist(PHI[,2], freq = FALSE, main='Gibbs draws', breaks=20,xlab=expression(gamma))
plot(cumsum(PHI[,2])/seq(1,S),type="l",main='Gibbs draws',
     xlab='Iteration number', ylab='cumulative mean of theta')
lines(seq(1,S),1*matrix(2.15,1,S),col="red",lwd=3)
acf(PHI[,2], main='Gibbs draws', lag.max = 20)
```



Last, for β

```
par(mfrow=c(2,2))
plot(PHI[,3], type="l", main='Gibbs draws',
     xlab="Iteration number", ylab=expression(beta))
hist(PHI[,3], freq = FALSE, main='Gibbs draws', breaks=50,xlab=expression(beta))
plot(cumsum(PHI[,3])/seq(1,S),type="l",main='Gibbs draws',
     xlab='Iteration number', ylab='cumulative mean of theta')
lines(seq(1,S),1*matrix(0.075,1,S),col="red",lwd=3)
acf(PHI[,3], main='Gibbs draws', lag.max = 20)
```



(c)

```
log_lambda <- log(PHI[,1])
```

The Gibbs sample mean of $log(\lambda)$ is

```
mean(log_lambda)
```

```
## [1] -0.1586736
```

The 95% sample credible interval of $log(\lambda)$ is

```
quantile(log_lambda, c(0.025, 0.975))
```

```
##       2.5%      97.5%
## -0.4558388  0.1267556
```

(d)

Let p the probability of $x_i = 0$. Consider that we have no information about p, since nothing is given from the question. We can set the prior distribution of p as

$$p \sim Beta(1,1) = Unif(0,1)$$

which reflects minimum information about the parameters.

since $n = 100, m = 50$ from the data, the **likelihood** and the **posterior** can be written as

$$x_i|p \sim Binom(100, \frac{1}{2})$$
$$p|x_i \sim Beta(51, 51)$$

It is known that the predictive distribution of x_i follows a **Beta-Binomial model**. If $\tilde{x}$ out of *m* samples are drawn, the predictive distribution is

$$p(\tilde{x}|x) = \binom{m}{\tilde{x}} \frac{\Gamma(1+1+100)}{\Gamma(1+50)\Gamma(1+100-50)} \frac{\Gamma(1+50+\tilde{x})\Gamma(1+100-m-50-\tilde{x})}{\Gamma(1+1+100+m)}$$
$$p(\tilde{x}|x) = \binom{m}{\tilde{x}} \frac{\Gamma(102)}{\Gamma(51)\Gamma(51)} \frac{\Gamma(51+\tilde{x})\Gamma(51-m-\tilde{x})}{\Gamma(102+m)}$$

Question 4

(a)

$$

$$X_i|\mu, \sigma^2 \sim N(\mu, \sigma^2)$$
$$\theta_1 \sim N(\mu_1, \sigma_1^2)$$
$$\theta_2 \sim N(\mu_2, \sigma_2^2)$$

$$Thenthejointposteriordistributioncanbecalculatedby$$

_1, 2, ^2 | *X_i* {i=1}^{n} exp(-(y_i - _1 - _2)^2) exp(-(_1 - _1)^2) exp(-(_2 - _2)^2)

$$The fullconditionaldistributionof\$\theta_1\$is$$

_1,| *2, X_i* {i=1}^{n} exp(-(x_i - _1 - _2)^2) exp(-(_1 - _1)^2)

$$Since\$\tau_1^2 = \frac{1+n\sigma_1^2}{n}\$and\$\mu_1 = \tau_1^2(\frac{\mu_1}{\sigma_1^2} + \frac{n\bar{x}}{1})\$$$

_1,| _2, X_i N(_1, _1^2) = N(_1^2(+),) $$

The full conditional distribution of θ_2 is

$$\theta_2, |\theta_1, X_i \propto \prod_{i=1}^{n} \frac{1}{\sigma} exp(-\frac{1}{2\sigma^2}(x_i - \theta_1 - \theta_2)^2) \times \frac{1}{\sigma_2} exp(-\frac{1}{2\sigma_2^2}(\theta_2 - \mu_2)^2)$$

And by the same way,

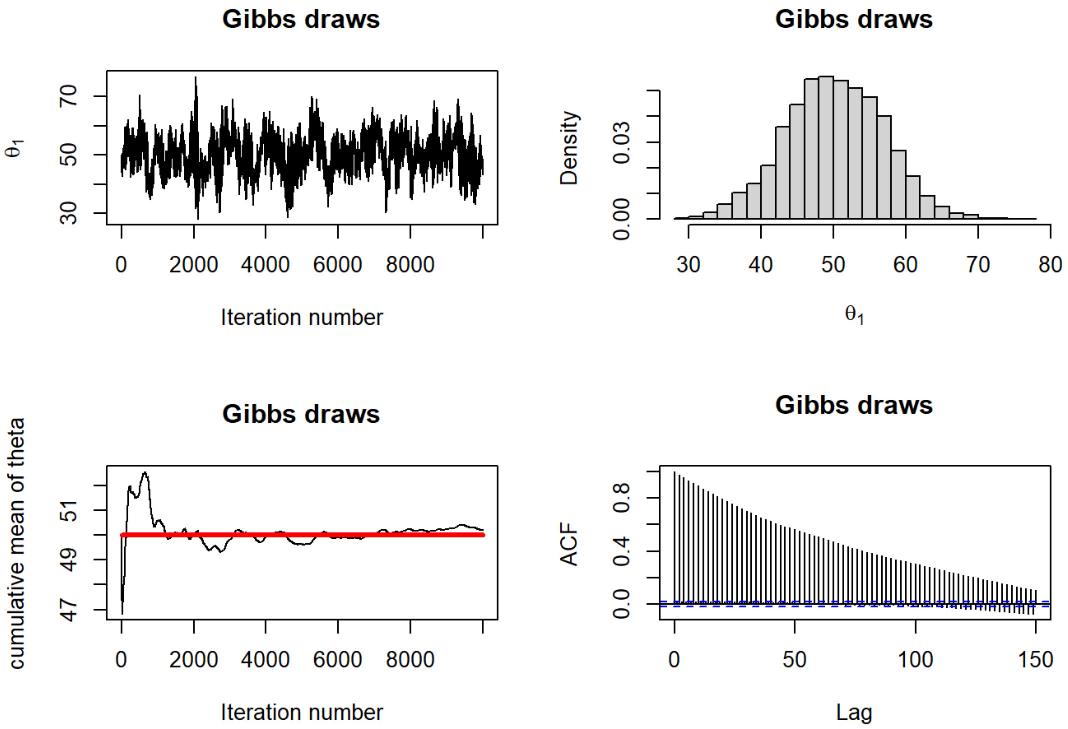
$$\theta_2, |\theta_1, X_i \sim N(\mu_2, \tau_2^2) = N(\tau_2^2(\frac{\mu_2}{\sigma_1^2} + \frac{n\bar{x}}{1}), \frac{\sigma_2^2}{1+n\sigma_2^2})$$

(b)

```
mu1<-50; mu2<-50; sigma1<-10; sigma2<-10;
x<-0;
tau1 <- sigma1^2/(1+sigma1^2); tau2 <-sigma2^2/(1+sigma2^2)
S<-10^4
PHI<-matrix(nrow=S,ncol=3)
PHI[1,1] <- mu1
PHI[1,2] <- mu2
PHI[1,3] <- PHI[1,1]+PHI[1,2]
### Gibbs sampling
for(s in 2:S) {
  PHI[s,1]<- rnorm(1, tau1*(mu1/(sigma1^2)+PHI[s-1,2]), tau1)
  PHI[s,2]<- rnorm(1, tau1*(mu2/(sigma2^2)+PHI[s-1,1]), tau2)
  PHI[s,3] <- PHI[s,1]+PHI[s,2]
}
```

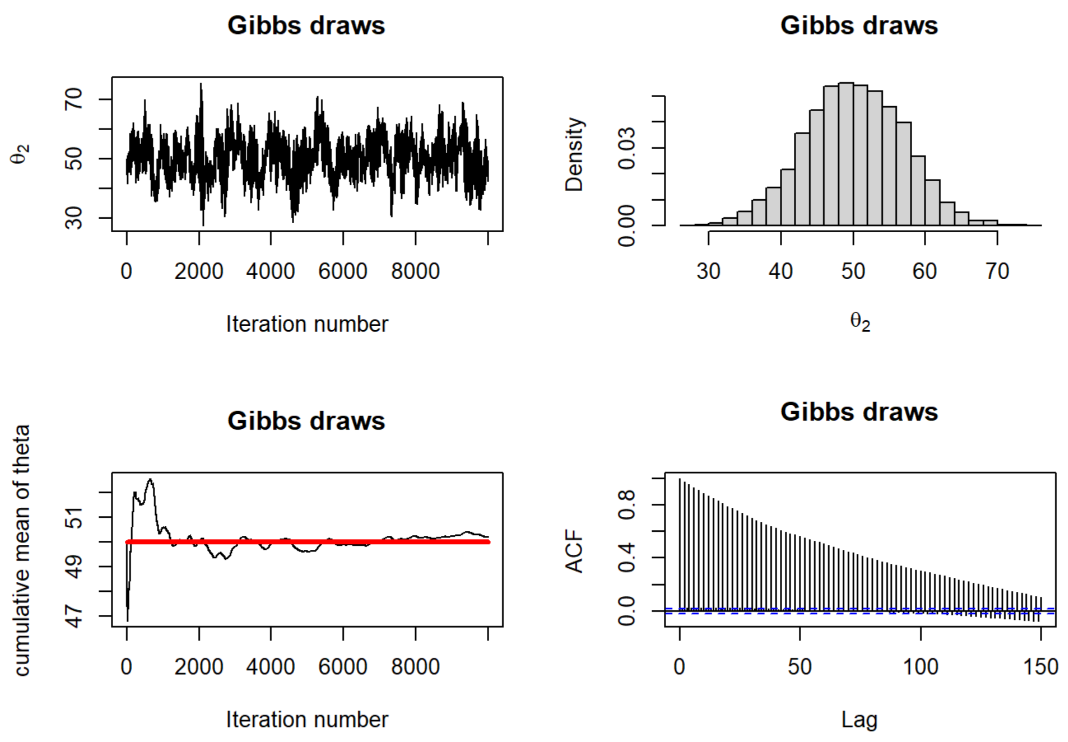
Now the Gibbs samples of $\theta_1, \theta_2, \theta_1 + \theta_2$ are made. To see whether they are mixed well, we will check several graphs including ACF. First, for the θ_1,

```
par(mfrow=c(2,2))
plot(PHI[,1], type="l", main='Gibbs draws',
     xlab="Iteration number", ylab=expression(theta[1]))
hist(PHI[,1], freq = FALSE, main='Gibbs draws', breaks=20,xlab=expression(theta[1]))
plot(cumsum(PHI[,1])/seq(1,S),type="l",main='Gibbs draws',
     xlab='Iteration number', ylab='cumulative mean of theta')
lines(seq(1,S),1*matrix(50,1,S),col="red",lwd=3)
acf(PHI[,1], main='Gibbs draws', lag.max = 150)
```



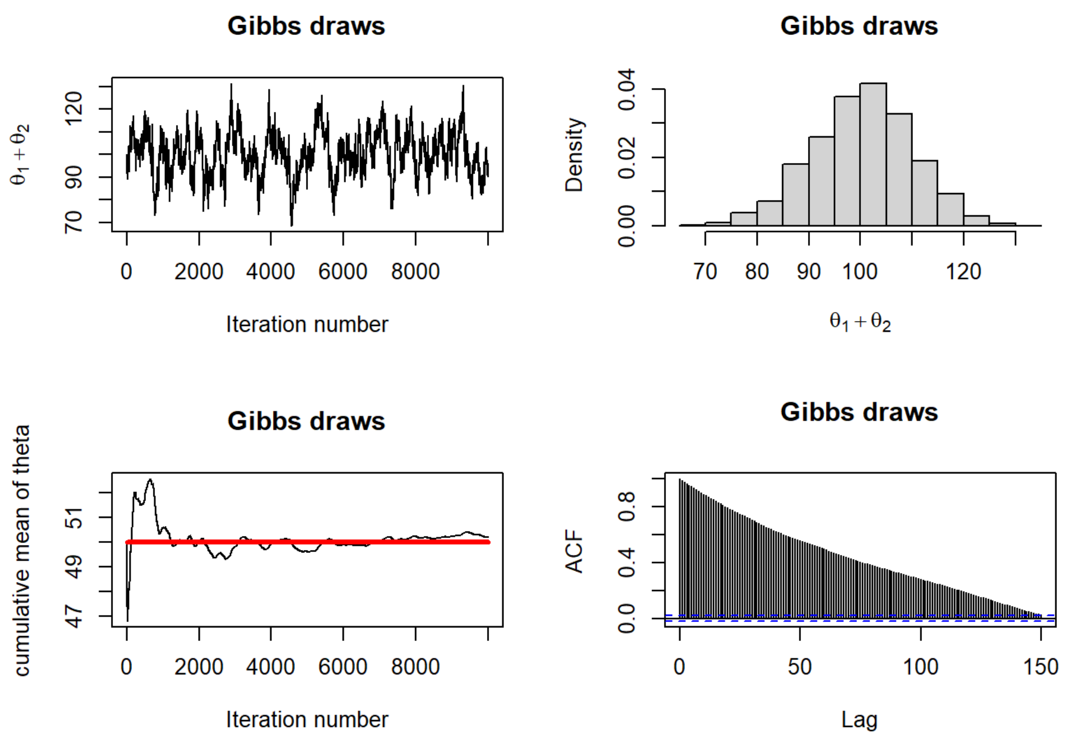
For θ_2

```
par(mfrow=c(2,2))
plot(PHI[,2], type="l", main='Gibbs draws',
     xlab="Iteration number", ylab=expression(theta[2]))
hist(PHI[,2], freq = FALSE, main='Gibbs draws', breaks=20,xlab=expression(theta[2]))
plot(cumsum(PHI[,2])/seq(1,S),type="l",main='Gibbs draws',
     xlab='Iteration number', ylab='cumulative mean of theta')
lines(seq(1,S),1*matrix(50,1,S),col="red",lwd=3)
acf(PHI[,2], main='Gibbs draws', lag.max = 150)
```



Last, for $\theta_1 + \theta_2$

```
par(mfrow=c(2,2))
plot(PHI[,3], type="l", main='Gibbs draws',
     xlab="Iteration number", ylab=expression(theta[1]+theta[2]))
hist(PHI[,3], freq = FALSE, main='Gibbs draws', breaks=20,xlab=expression(theta[1]+theta[2]))
plot(cumsum(PHI[,2])/seq(1,S),type="l",main='Gibbs draws',
     xlab='Iteration number', ylab='cumulative mean of theta')
lines(seq(1,S),1*matrix(50,1,S),col="red",lwd=3)
acf(PHI[,3], main='Gibbs draws', lag.max = 150)
```

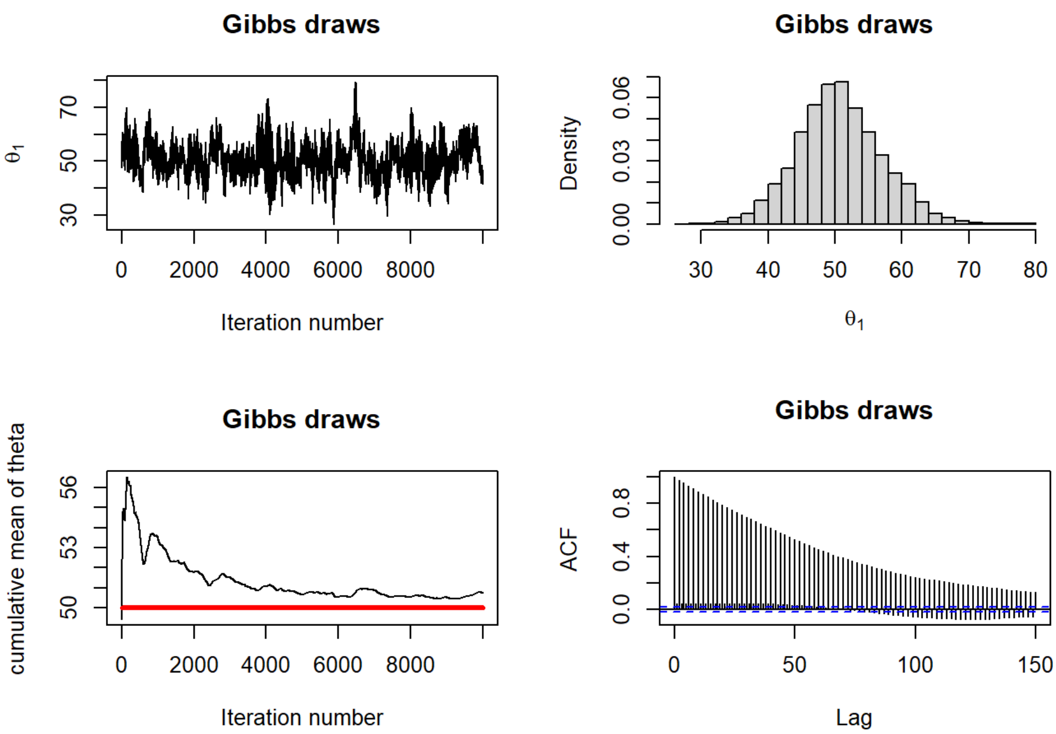


(c)

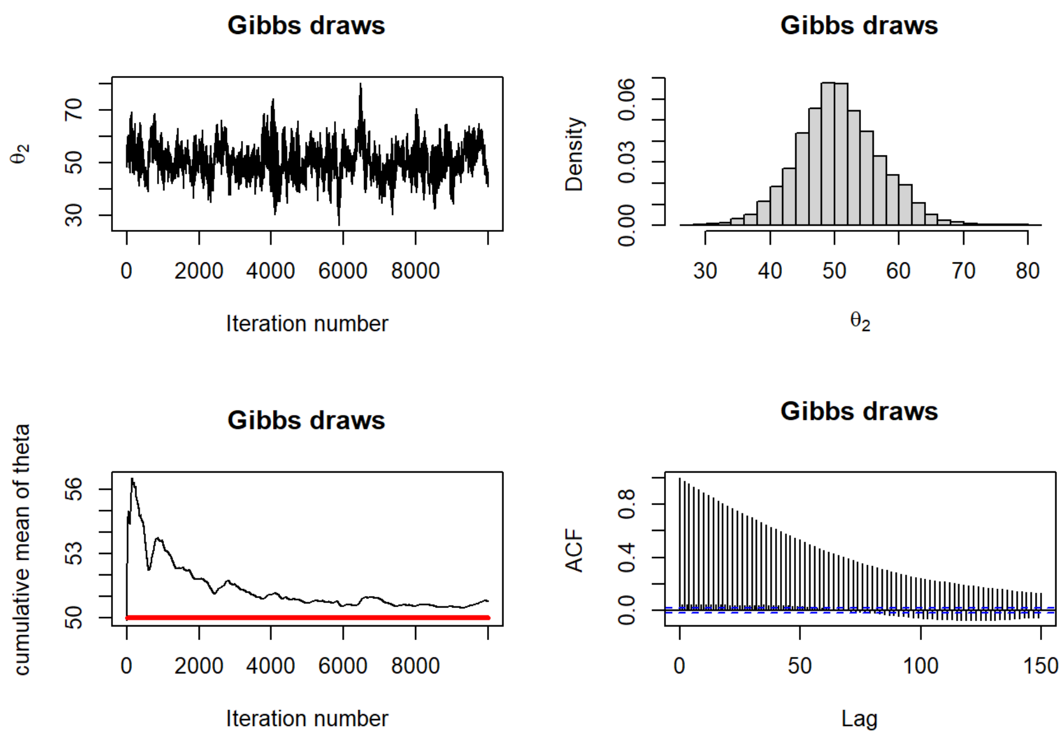
```
mu1<-50; mu2<-50; sigma1<-10; sigma2<-10;
x<-0;
tau1 <- sigma1^2/(1+sigma1^2); tau2 <-sigma2^2/(1+sigma2^2)
S<-10^4
PHI<-matrix(nrow=S,ncol=3)
PHI[1,1] <- mu1
PHI[1,2] <- mu2
PHI[1,3] <- PHI[1,1]+PHI[1,2]
### Gibbs sampling
for(s in 2:S) {
  PHI[s,1]<- rnorm(1, tau1*(mu1/(sigma1^2)+PHI[s-1,2]), tau1)
  PHI[s,2]<- rnorm(1, tau1*(mu2/(sigma2^2)+PHI[s-1,1]), tau2)
  PHI[s,3] <- PHI[s,1]+PHI[s,2]
}
```

Now the Gibbs samples of $\theta_1, \theta_2, \theta_1 + \theta_2$ are made. To see whether they are mixed well, we will check several graphs including ACF.

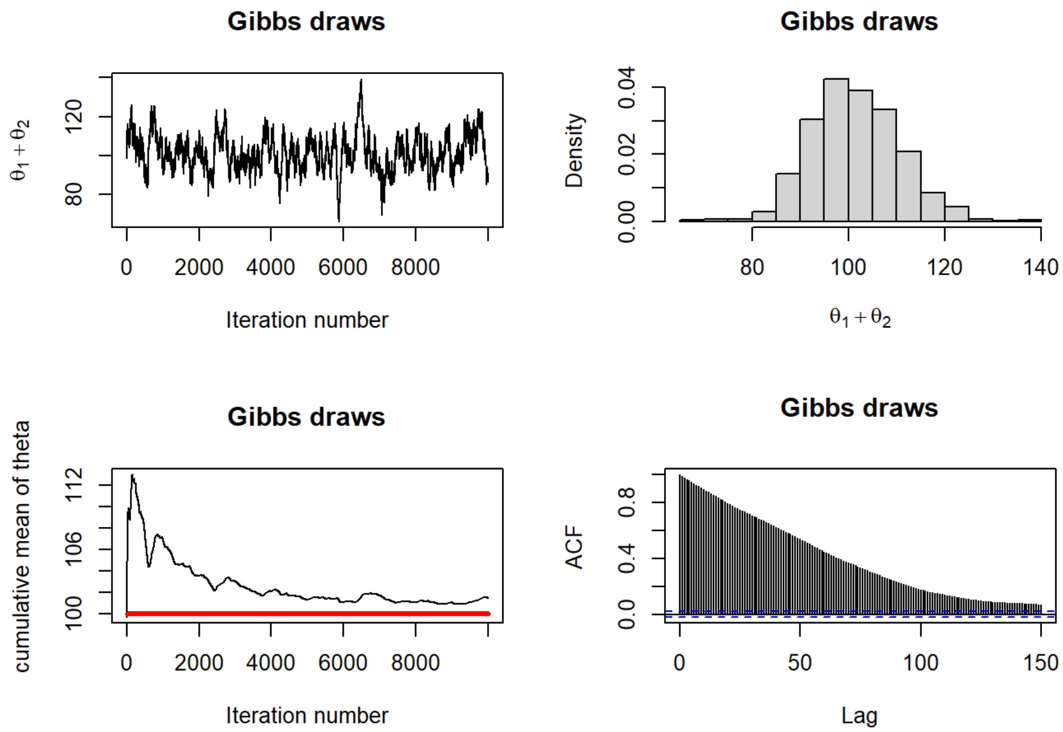
```
par(mfrow=c(2,2))
plot(PHI[,1], type="l", main='Gibbs draws',
     xlab="Iteration number", ylab=expression(theta[1]))
hist(PHI[,1], freq = FALSE, main='Gibbs draws', breaks=20,xlab=expression(theta[1]))
plot(cumsum(PHI[,1])/seq(1,S),type="l",main='Gibbs draws',
     xlab='Iteration number', ylab='cumulative mean of theta')
lines(seq(1,S),1*matrix(50,1,S),col="red",lwd=3)
acf(PHI[,1], main='Gibbs draws', lag.max = 150)
```




```
par(mfrow=c(2,2))
plot(PHI[,2], type="l", main='Gibbs draws',
     xlab="Iteration number", ylab=expression(theta[2]))
hist(PHI[,2], freq = FALSE, main='Gibbs draws', breaks=20,xlab=expression(theta[2]))
plot(cumsum(PHI[,2])/seq(1,S),type="l",main='Gibbs draws',
     xlab='Iteration number', ylab='cumulative mean of theta')
lines(seq(1,S),1*matrix(50,1,S),col="red",lwd=3)
acf(PHI[,2], main='Gibbs draws', lag.max = 150)
```




```
par(mfrow=c(2,2))
plot(PHI[,3], type="l", main='Gibbs draws',
     xlab="Iteration number", ylab=expression(theta[1]+theta[2]))
hist(PHI[,3], freq = FALSE, main='Gibbs draws', breaks=20,xlab=expression(theta[1]+theta[2]))
plot(cumsum(PHI[,3])/seq(1,S),type="l",main='Gibbs draws',
     xlab='Iteration number', ylab='cumulative mean of theta')
lines(seq(1,S),1*matrix(100,1,S),col="red",lwd=3)
acf(PHI[,3], main='Gibbs draws', lag.max = 150)
```



Compared to the previous Gibbs samples in **(b)**, the samples do not converge to the target μ values. This is because the prior variance was too large. Since Gibbs samplers converges slower than MC samples, large prior variance led to this spurious effect.

Question 5

(a)

I will use a famous dataset named **Galton's Height** dataset. This is a dataset of father's height and child's hight in inches. I will reproduce the units from **inches** to **centimeters**.

```
library(HistData)
data("Galton")
Galton.cm <- Galton*2.54; Galton.cm <- round(Galton.cm, 1); head(Galton.cm)
```

```
##   parent child
## 1  179.1 156.7
## 2  174.0 156.7
## 3  166.4 156.7
## 4  163.8 156.7
## 5  162.6 156.7
## 6  171.4 158.0
```

```
head(Galton.cm)
```


```
##   parent child
## 1  179.1 156.7
## 2  174.0 156.7
## 3  166.4 156.7
## 4  163.8 156.7
## 5  162.6 156.7
## 6  171.4 158.0
```


(b)

The data looks like above, and the scatter plot of this data is

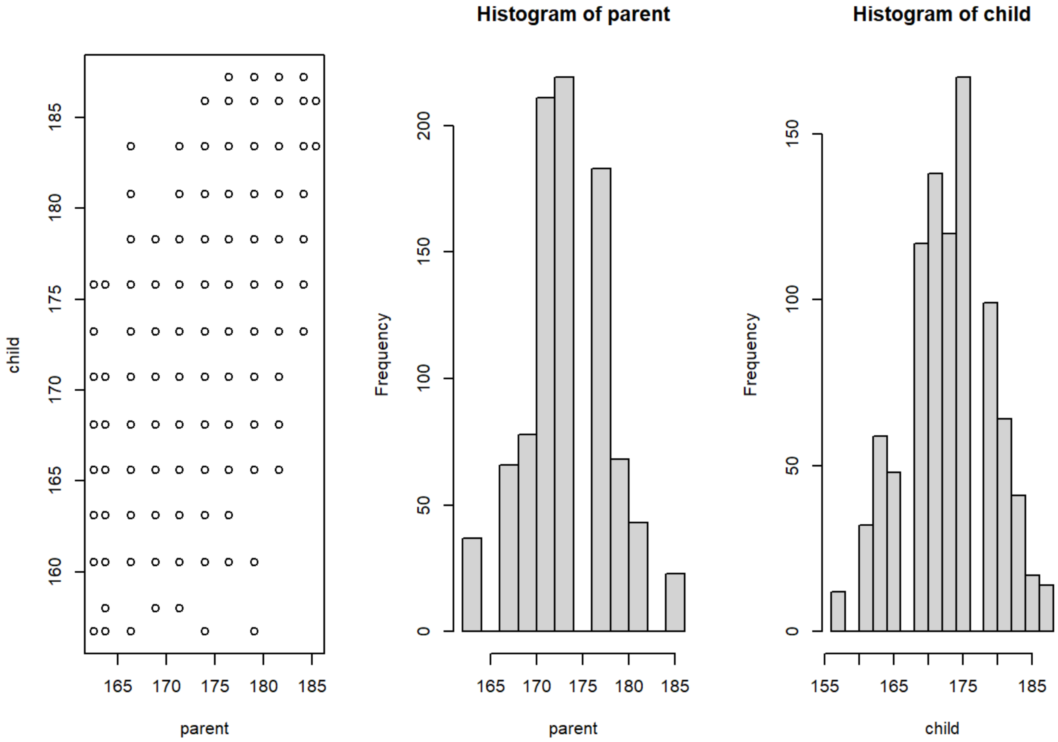
```
library(dplyr)
```

```
##
## 다음의 패키지를 부착합니다: 'dplyr'
```

```
## The following objects are masked from 'package:stats':
##
##     filter, lag
```

```
## The following objects are masked from 'package:base':
##
##     intersect, setdiff, setequal, union
```

```
attach(Galton.cm)
par(mfrow=c(1,3))
plot(parent, child)
hist(parent)
hist(child)
```



Some summary statistics about this data are

```
mean(parent); mean(child);
```

```
## [1] 173.4917
```

```
## [1] 172.9391
```

```
var(parent); var(child);
```

```
## [1] 20.63927
```

```
## [1] 40.91209
```

```
cor(parent, child)
```

```
## [1] 0.4590389
```

The mean of **parent** and **child** are quite similar, while variance are slightly different. The correlation is about 0.46, which means that it shows some linear relationship between those two variables.

(c)

According to the *ROK Military Manpower Administration* it is told that height data follows $N(173, 6^2)$. So I will set the prior distribution of precision as $\frac{1}{\sigma^2} \sim \Gamma(1, 36)$ where the prior mean of $\frac{1}{\sigma^2}$ is $\frac{1}{6^2}$. About the coefficients, I have no idea, but it looks like the correlation is similar to 0.5, so I will fit

$$\beta_0 \sim N(0, 1)$$
$$\beta_1 \sim N(0.5, 1)$$
$$\frac{1}{\sigma^2} \sim \Gamma(1, 36)$$

(d)

```
library(rjags)
```

```
## 필요한 패키지를 로딩중입니다: coda
```

```
## Linked to JAGS 4.3.1
```

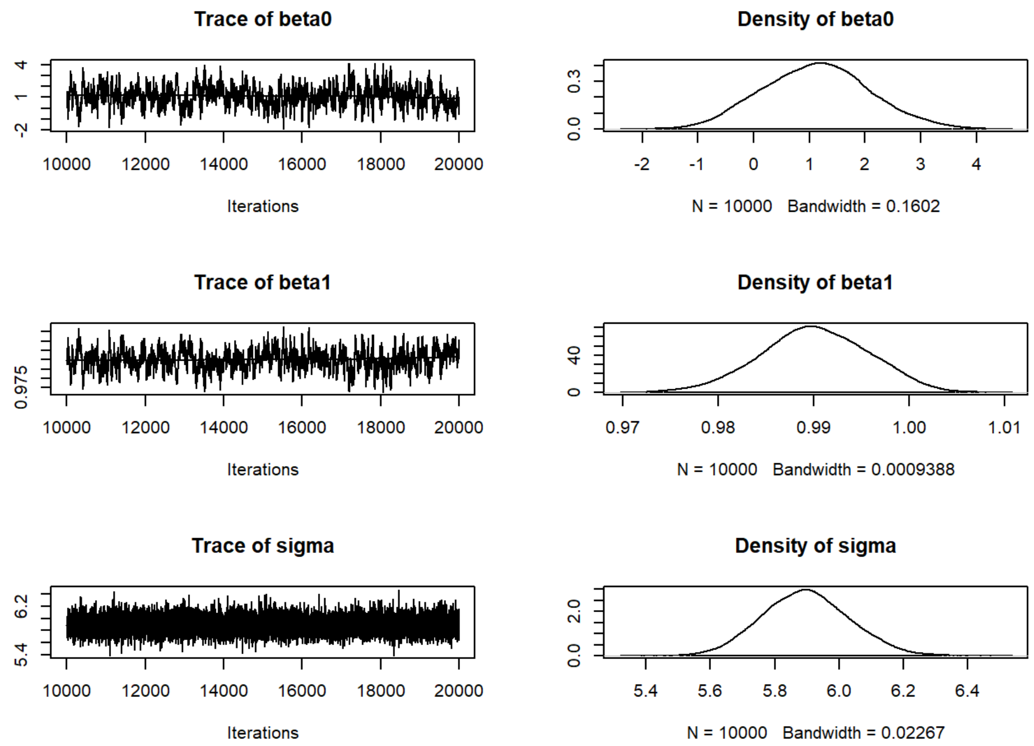
```
## Loaded modules: basemod,bugs
```

```
jags_code ="model{
for( i in 1:n) {
y[i] ~ dnorm(mu[i] , tau); mu[i] <- beta0+ beta1*x[i] }
beta0 ~ dnorm(0, 1); beta1 ~ dnorm(0.5, 1)
tau ~ dgamma(1, 36); sigma <- 1/sqrt(tau)}"
height = list(x=c(parent),
              y = c(child), n=nrow(Galton.cm))
jags_reg = jags.model(textConnection(jags_code), data=height)
```

```
## Compiling model graph
##    Resolving undeclared variables
##    Allocating nodes
## Graph information:
##    Observed stochastic nodes: 928
##    Unobserved stochastic nodes: 3
##    Total graph size: 1888
##
## Initializing model
```

```
update(jags_reg, 10000) #progress.bar="none")
samp <- coda.samples(jags_reg,variable.names=
                       c("beta0","beta1","sigma"), n.iter=10000)
summary(samp); plot(samp)
```

```
##
## Iterations = 10001:20000
## Thinning interval = 1
## Number of chains = 1
## Sample size per chain = 10000
##
## 1. Empirical mean and standard deviation for each variable,
##    plus standard error of the mean:
##
##         Mean       SD  Naive SE Time-series SE
## beta0 1.1241 0.953790 9.538e-03      0.0673469
## beta1 0.9901 0.005588 5.588e-05      0.0003928
## sigma 5.8967 0.134916 1.349e-03      0.0013492
##
## 2. Quantiles for each variable:
##
##          2.5%    25%    50%   75% 97.5%
## beta0 -0.6830 0.4690 1.1294 1.763 3.049
## beta1  0.9789 0.9864 0.9901 0.994 1.001
## sigma  5.6411 5.8037 5.8942 5.986 6.166
```




```
lm(height$y ~ height$x)
```

```
##
## Call:
## lm(formula = height$y ~ height$x)
##
## Coefficients:
## (Intercept)     height$x
##     60.8130       0.6463
```