STAT 431 Homework 5 Solutions

Robert Garrett

Problem 1

Part a

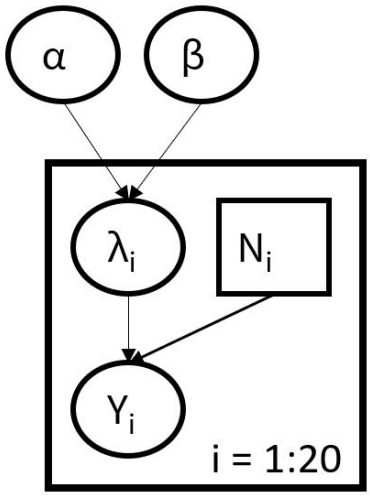


Figure 1: Problem 1 Model DAG

Part b

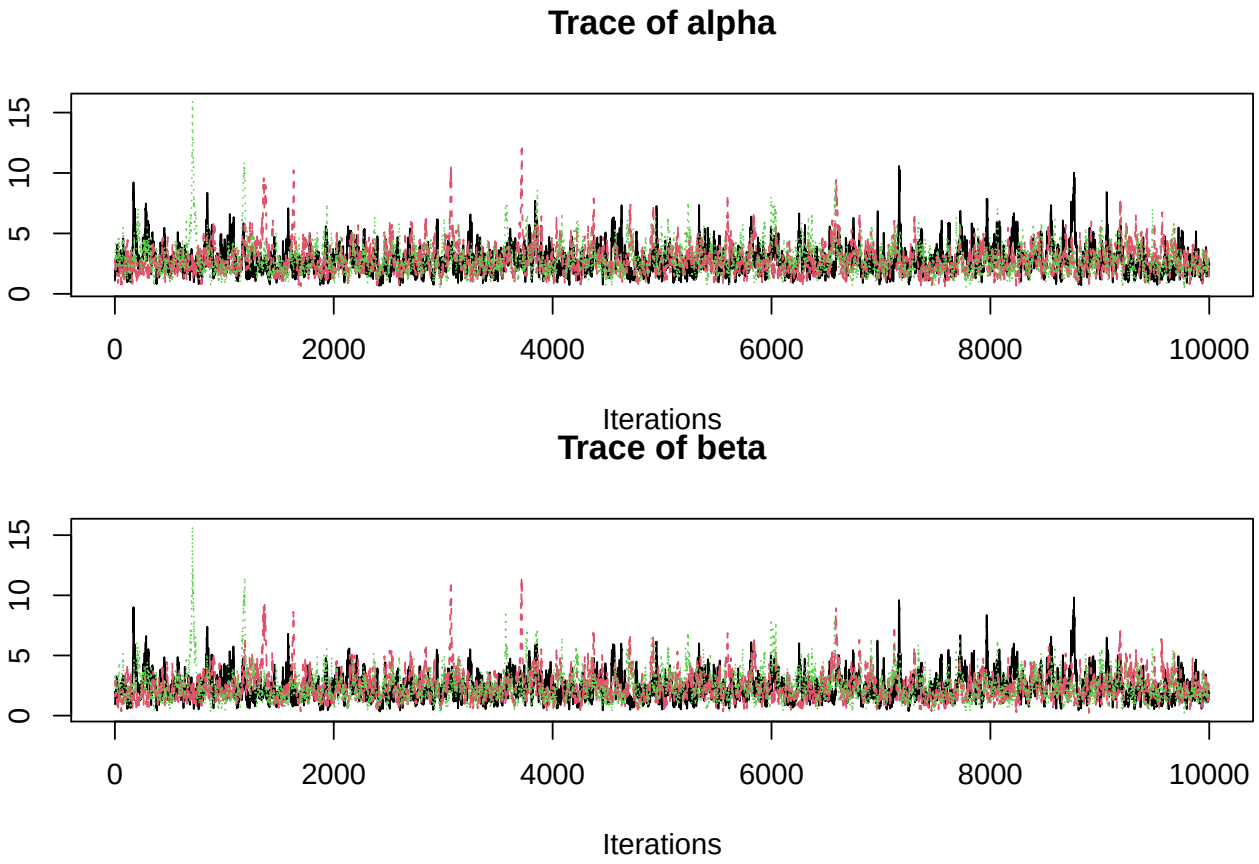
Initial values: The first chain was initialized with $\alpha = \beta = 1$, the second was initialized with $\alpha = \beta = 2$, and the last was initialized with $\alpha = \beta = 3$.

Confirming Convergence: I confirmed convergence using trace plots and plots of the Gelman statistics for α and β. The trace plots appeared sufficiently mixed after 1000 iterations or less. The Gelman statistic plots showed values converging to 1 within a margin of about 1/1000 after 5000 iterations or less.

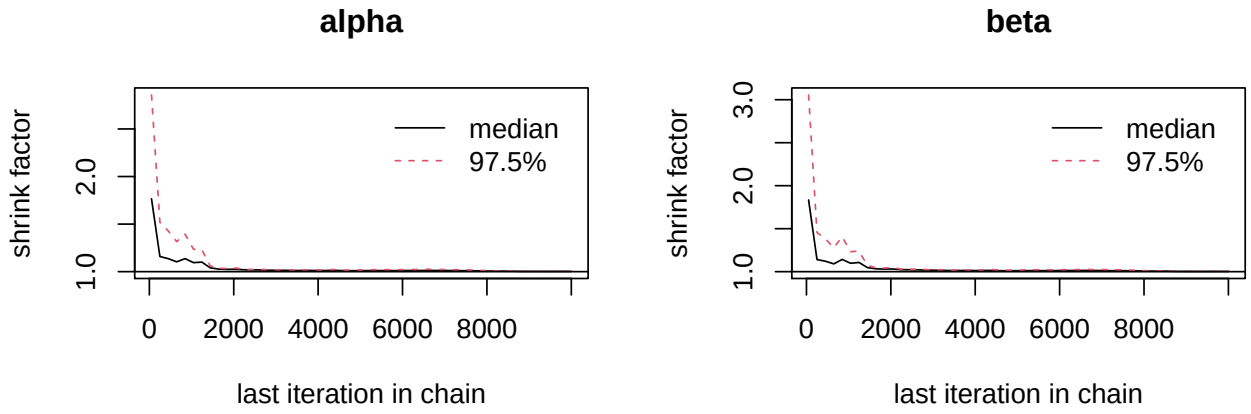
10,000 burn-in iterations: After 10,000 initial iterations, convergence was confirmed. So, I will burn these iterations and sample 100,000 more iterations on each chain.

```
data = read.table('airlinerdata.txt',header = T)
n = nrow(data)

d = list(y = data$y, N = data$N)
inits = list(list(alpha=1, beta=1),
             list(alpha=2, beta=2),
             list(alpha=3, beta=3))
m = jags.model("prob1model.bug", d, inits, n.chains = 3 ,n.adapt = 0)

samples =  coda.samples(m, c("alpha","beta"), n.iter = 10000)
```

```
traceplot(samples)
```




```
gelman.plot(samples, autoburnin=FALSE)
```



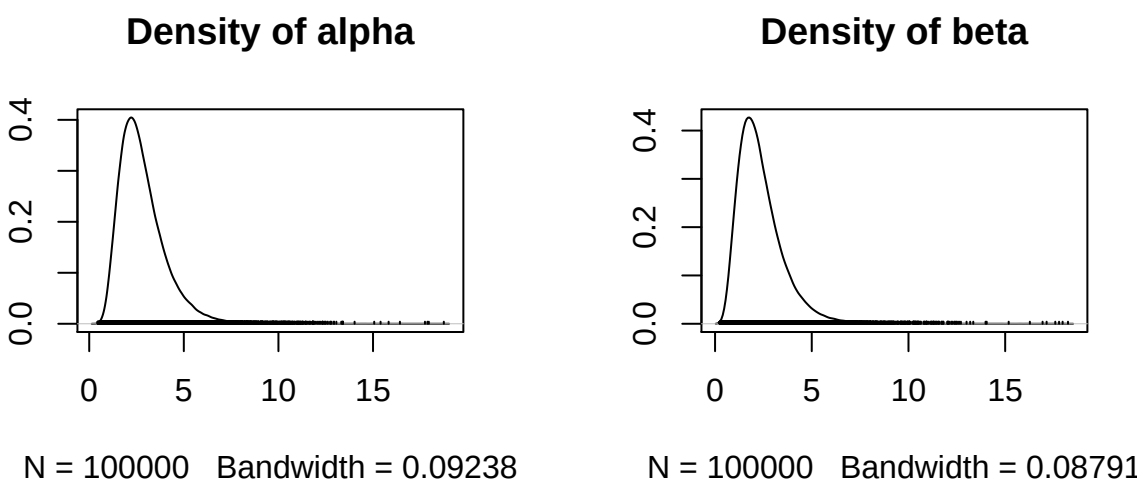
Part c

```
samples =  coda.samples(m, c("alpha","beta"), n.iter = 100000)
summary(samples)
```

```
##
## Iterations = 10001:110000
## Thinning interval = 1
## Number of chains = 3
```

```
## Sample size per chain = 1e+05
##
## 1. Empirical mean and standard deviation for each variable,
##    plus standard error of the mean:
##
##        Mean    SD Naive SE Time-series SE
## alpha 2.810 1.233 0.002251        0.01206
## beta  2.357 1.182 0.002158        0.01155
##
## 2. Quantiles for each variable:
##
##         2.5%   25%   50%   75% 97.5%
## alpha 1.1247 1.950 2.577 3.405 5.821
## beta  0.7762 1.534 2.128 2.918 5.263
```

```
plot(samples, trace=F)
```



Problem 2

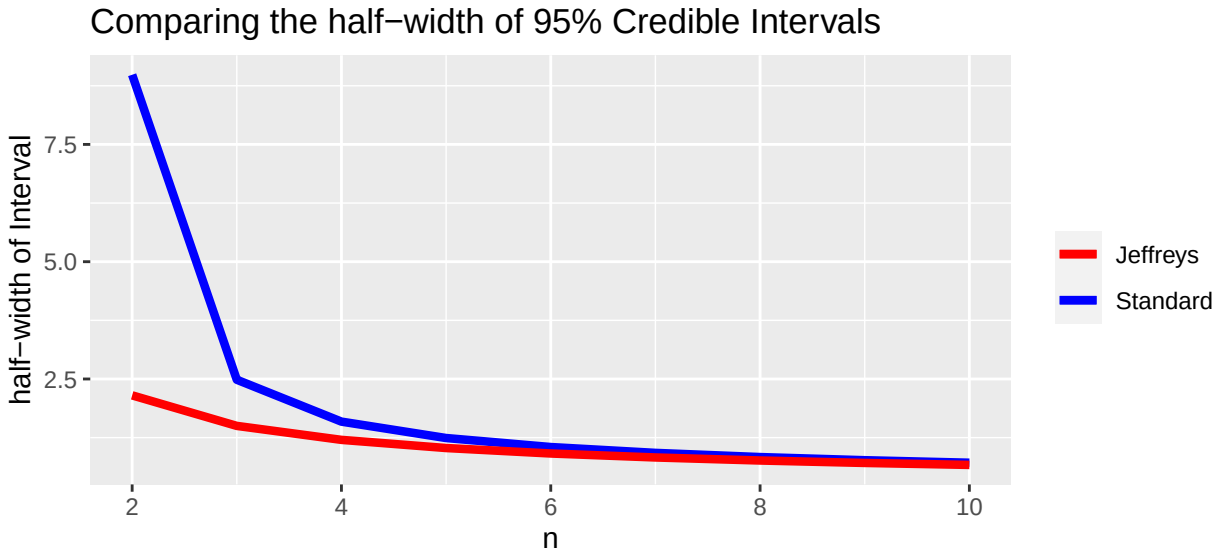
Part a

95% Credible interval for the standard noninformative prior: $\bar{y} \pm t_{0.025,n-1}\sqrt{\frac{s^2}{n}}$

95% Credible interval for the Jeffreys prior: $\bar{y} \pm t_{0.025,n}\frac{\sqrt{(n-1)s^2}}{n}$

Part b

```
n = 2:10
half_width_standard = function(n){
  qt(0.975,n-1)/sqrt(n)
}
half_width_Jeffreys = function(n){
  qt(0.975,n)*sqrt(n-1)/n
}
ggplot()+
  geom_line(aes(n,half_width_standard(n)),color='Standard'),size=1.5)+
  geom_line(aes(n,half_width_Jeffreys(n)),color='Jeffreys'),size=1.5)+
  scale_color_manual(values = c('red','blue'))+
  labs(y='half-width of Interval',color=NULL,
       title='Comparing the half-width of 95% Credible Intervals')
```



The Standard prior appears to produce wider intervals than the Jeffreys prior.

Problem 3

JAGS Code:

```
model {
  for (i in 1:length(y)) {
    y[i] ~ dpois(N[i] * lambda[i])
    lambda[i] ~ dgamma(alpha, beta)
  }

  lambda_new ~ dgamma(alpha, beta)

  alpha ~ dgamma(0.001, 0.001)
  beta ~ dgamma(0.001, 0.001)
}
```

Part b

First, initialize the sampler with the same 10,000 burn-in as before:

```
# same data and initial values as before
m = jags.model("prob3model.bug", d, inits, n.chains=3)
samples =  coda.samples(m, "lambdanew", n.iter=10000)
```

Now, run the sampler for the 100,000 iterations and get our estimates:

```
samples = coda.samples(m, "lambdanew", n.iter=100000)
summary(samples)
```

```
##
## Iterations = 11001:111000
## Thinning interval = 1
## Number of chains = 3
## Sample size per chain = 1e+05
##
## 1. Empirical mean and standard deviation for each variable,
##    plus standard error of the mean:
##
##           Mean             SD       Naive SE Time-series SE
##       1.248079       0.877010       0.001601       0.001731
##
## 2. Quantiles for each variable:
##
##   2.5%    25%    50%    75%  97.5%
## 0.1682 0.6504 1.0604 1.6198 3.4479
```

Our posterior mean for λ_{new} is approimately 1.248.

Our posterior standard deviation is approximately 0.877.

Our 95% credible interval based on posterior samples is approximately $(0.168, 3.448)$