MTH422_Assignment-4

Charitha

2024-03-28

Assignment - 4

question - 1

(10)

Given the data is the number of marine bivalve species discovered each year from $2010 - 2015$ was $64, 13, 33, 18, 30, 20$.

Given Y_t as the number of species discovered in year $2009 + t$.

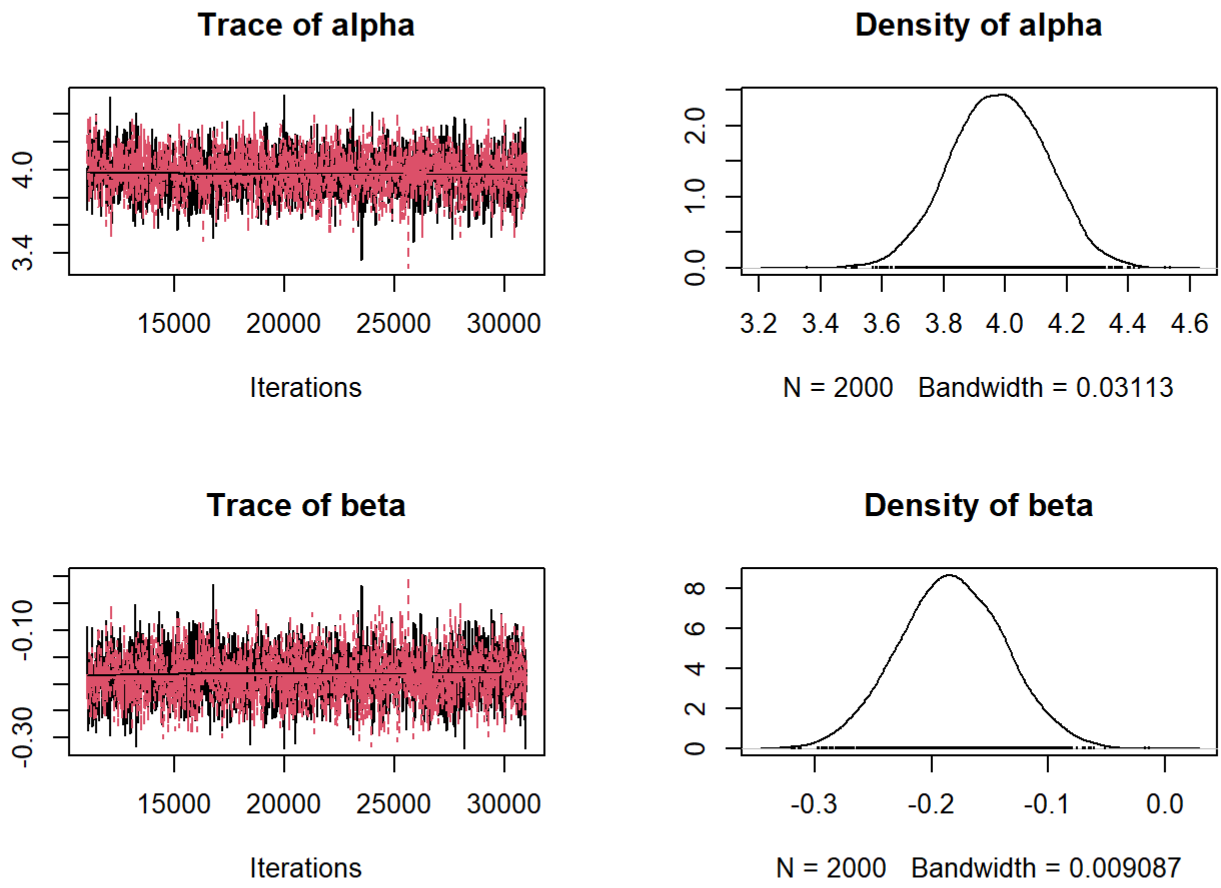
$$\begin{aligned} Y_t|\alpha, \beta &\sim \text{Poisson}(\lambda_t) \\ \lambda_t &= \exp(\alpha + \beta t) \\ \alpha, \beta &\sim \text{Normal}(0, 100) \end{aligned}$$

Using the above information, summarize the posterior of α and β by using JAGS to fit the model.

```
## Loading required package: coda
```

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

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



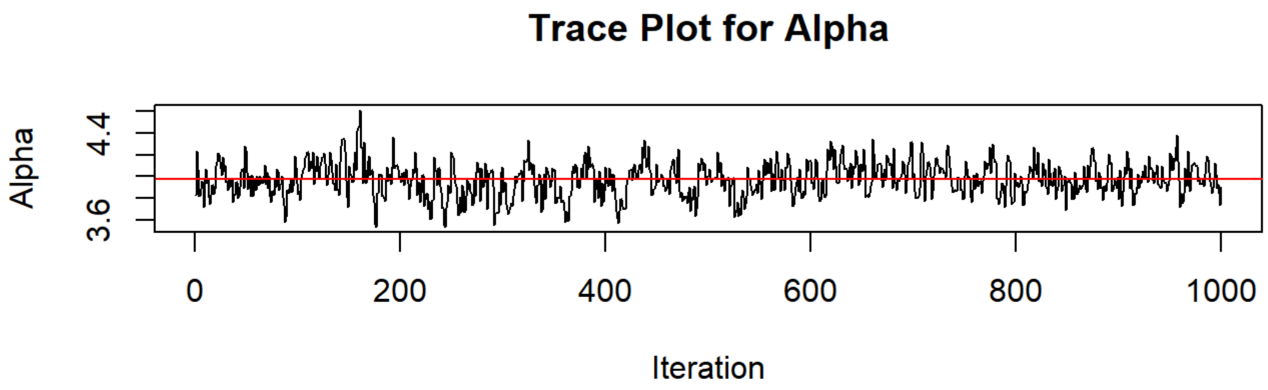

```
## Summary of Samples
```

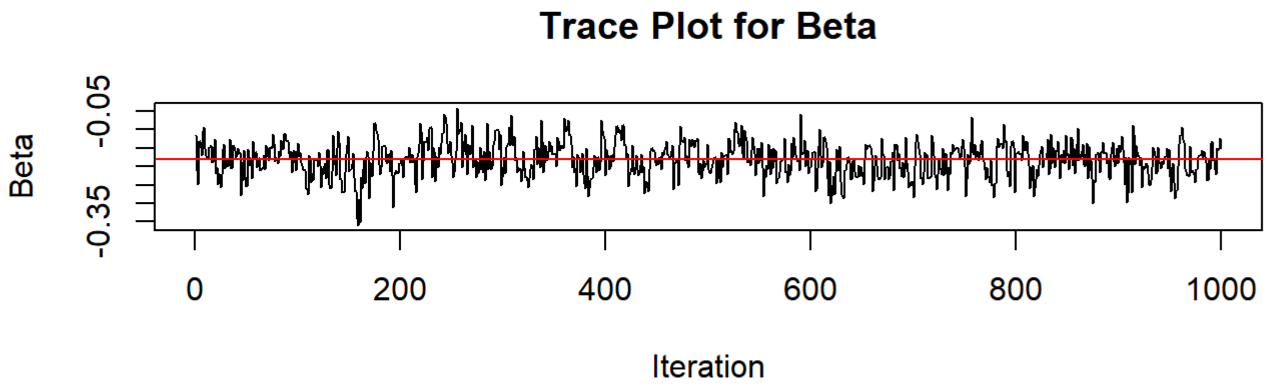
```
##
## Iterations = 11010:31000
## Thinning interval = 10
## Number of chains = 2
## Sample size per chain = 2000
##
## 1. Empirical mean and standard deviation for each variable,
##    plus standard error of the mean:
##
##          Mean      SD Naive SE Time-series SE
## alpha  3.9780 0.15425 0.002439      0.0030212
## beta  -0.1826 0.04503 0.000712      0.0008625
##
## 2. Quantiles for each variable:
##
##          2.5%     25%     50%     75%    97.5%
## alpha  3.6745  3.8730  3.9785  4.0856  4.27131
## beta  -0.2698 -0.2129 -0.1826 -0.1519 -0.09314
```

As the β (slope) is significantly different from zero, it suggests that the rate of discovery is changing over time.We can also look at the 95% credible intervals to see if they include zero, which would indicate no significant change.

(11)

Solving the above in **Metropolis sampler**






```
## Acceptance ratio of alpha is 0.2394
```

```
## Acceptance ratio of beta is 0.2365
```

```
## Summary of samples using Metropolis sampler
```

```
##        V1               V2
##  Min.   :3.540   Min.   :-0.36014
##  1st Qu.:3.872   1st Qu.:-0.21205
##  Median :3.973   Median :-0.17944
##  Mean   :3.973   Mean   :-0.18141
##  3rd Qu.:4.076   3rd Qu.:-0.15030
##  Max.   :4.609   Max.   :-0.04075
```

question - 2

Normal mixture model

$$Y_i|\theta \sim f(y|\theta)$$
$$f(y|\theta) = \frac{1}{2}[\phi(y-\theta) + \phi(y)]$$
$$\phi(z) = \frac{1}{\sqrt{2\pi}}\exp(\frac{-z^2}{2})$$

(a)

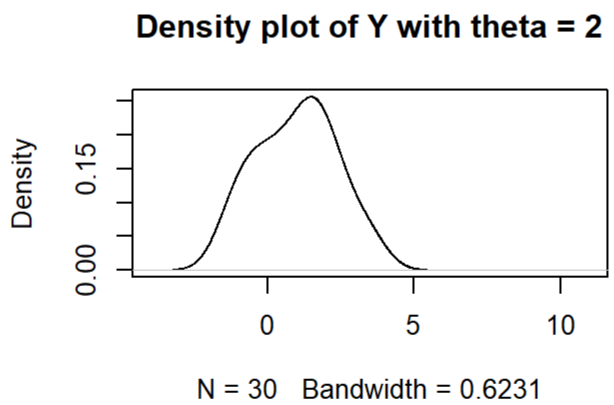
```
rm(list = ls())
set.seed(27695)
theta_true <- 4
n <- 30
B <- rbinom(n,1,0.5)
Y <- rnorm(n,B*theta_true,1)
```

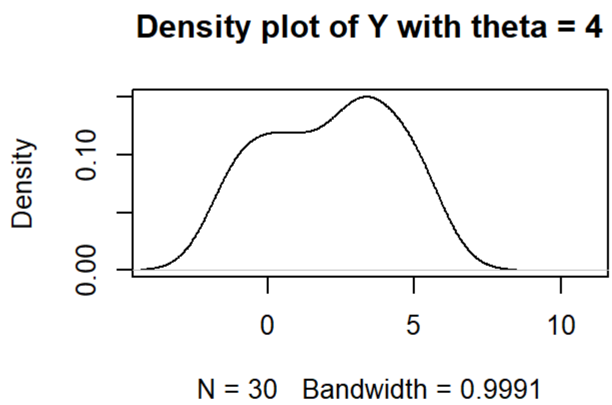
The above R code generates samples from $f(y|\theta)$.

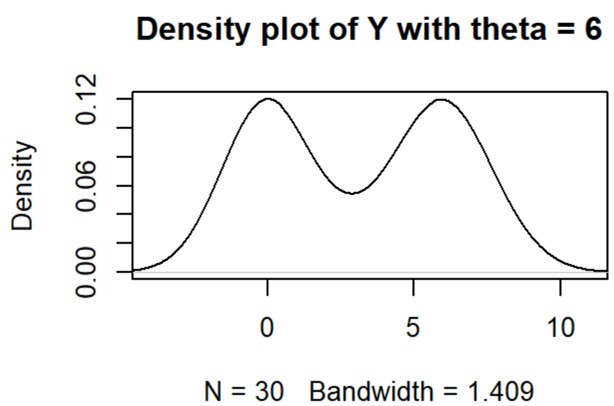
- **B** gives a vector of length n containing binary indicators $(0, 1)$ with equal probability.
- **Y** gives a vector of samples from mixture of two normal distributions.
  - If B[i] = 0, then Y[i] is generated from the density function represents the standard normal distribution N(0,1).
  - If B[i] = 1, the density function represents the shifted normal distribution N$(\theta_{true}, 1)$.

(b)

Plot $f(y|\theta)$ for $y \in [-3, 10]$ separately for $\theta = \{2, 4, 6\}$.







(c)

Assume prior of $\theta \sim$ Normal$(0, 10^2)$. MAP estimator of θ

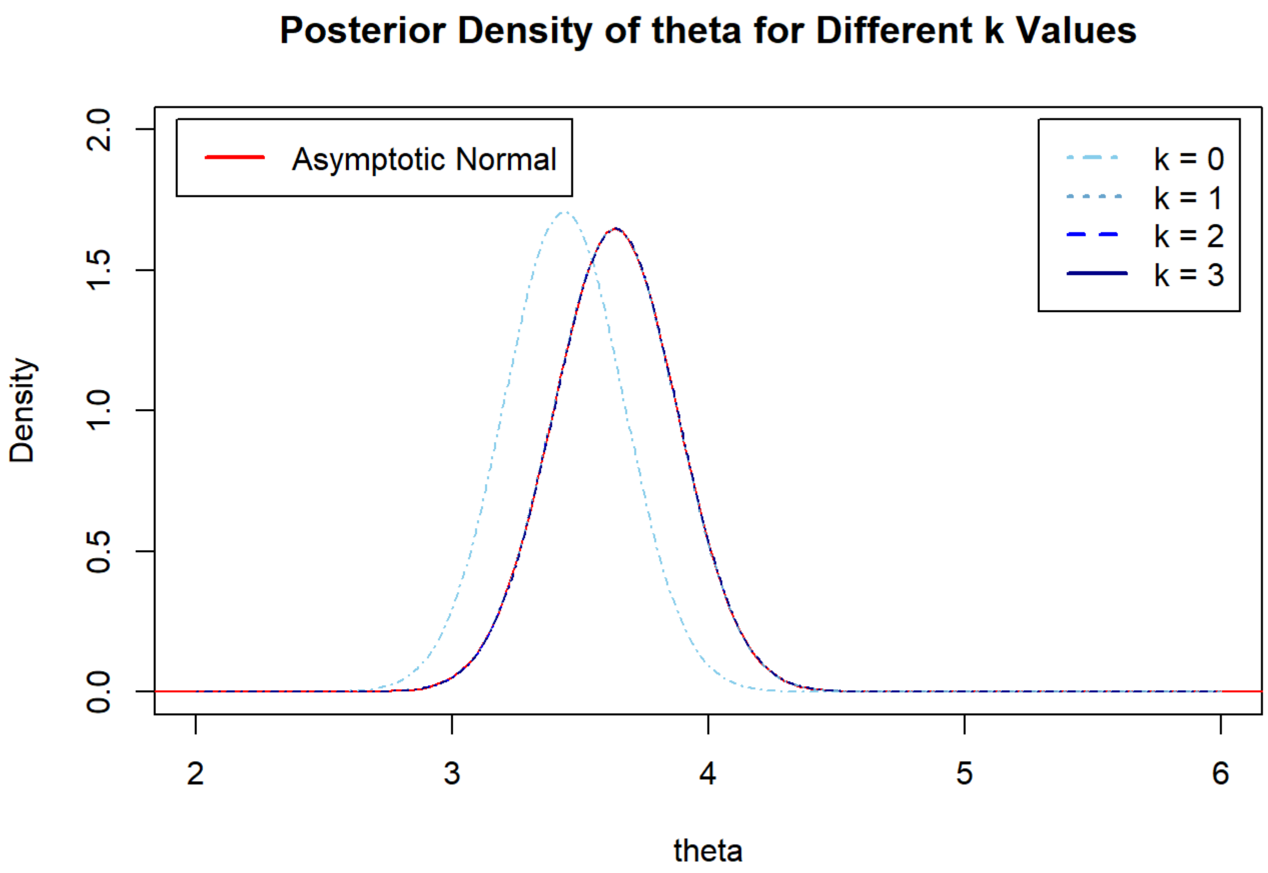
```
## MAP estimator of theta is  3.63574
```

```
## Standard deviation is  0.2420577
```

(d)

Suppose the prior distribution is $\theta \sim N(0, 10^k)$.

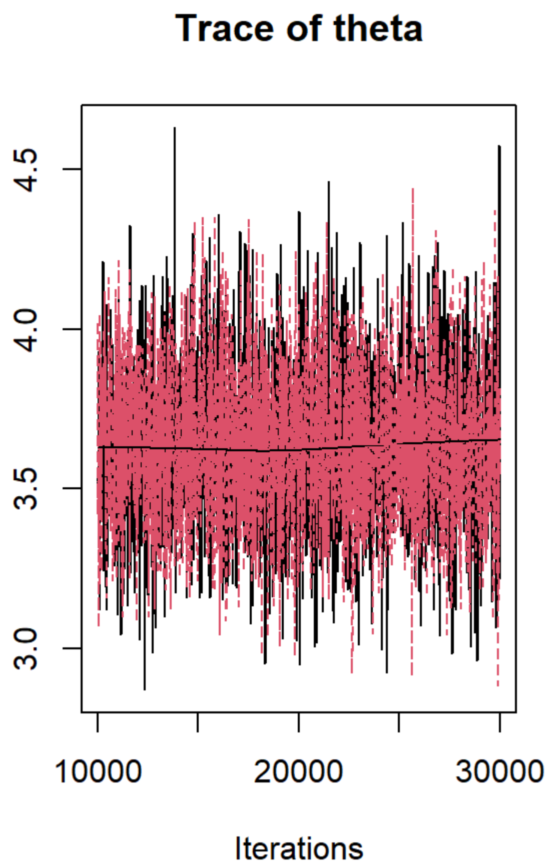
$k \in \{0, 1, 2, 3\}$

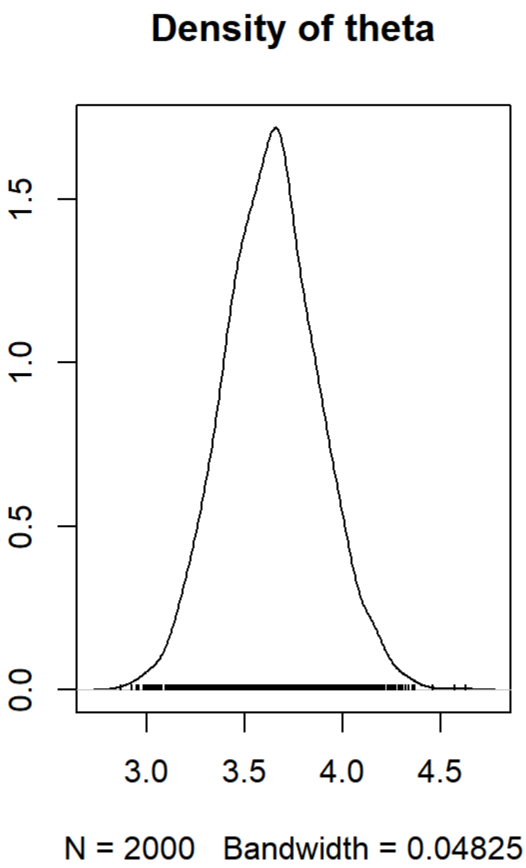



```
##   k_values map_est_k     sd_k
## 1        0  3.437913 0.2335890
## 2        1  3.635740 0.2420577
## 3        2  3.637850 0.2421500
## 4        3  3.637871 0.2421509
```

(e)

Use JAGS to fit this model via its mixture representation $Y_i|B_i, \theta \sim Normal(B_i\theta, 1)$, where $B_i \sim Bernoulli(0.5)$ and $\theta \sim Normal(0, 10^2)$.

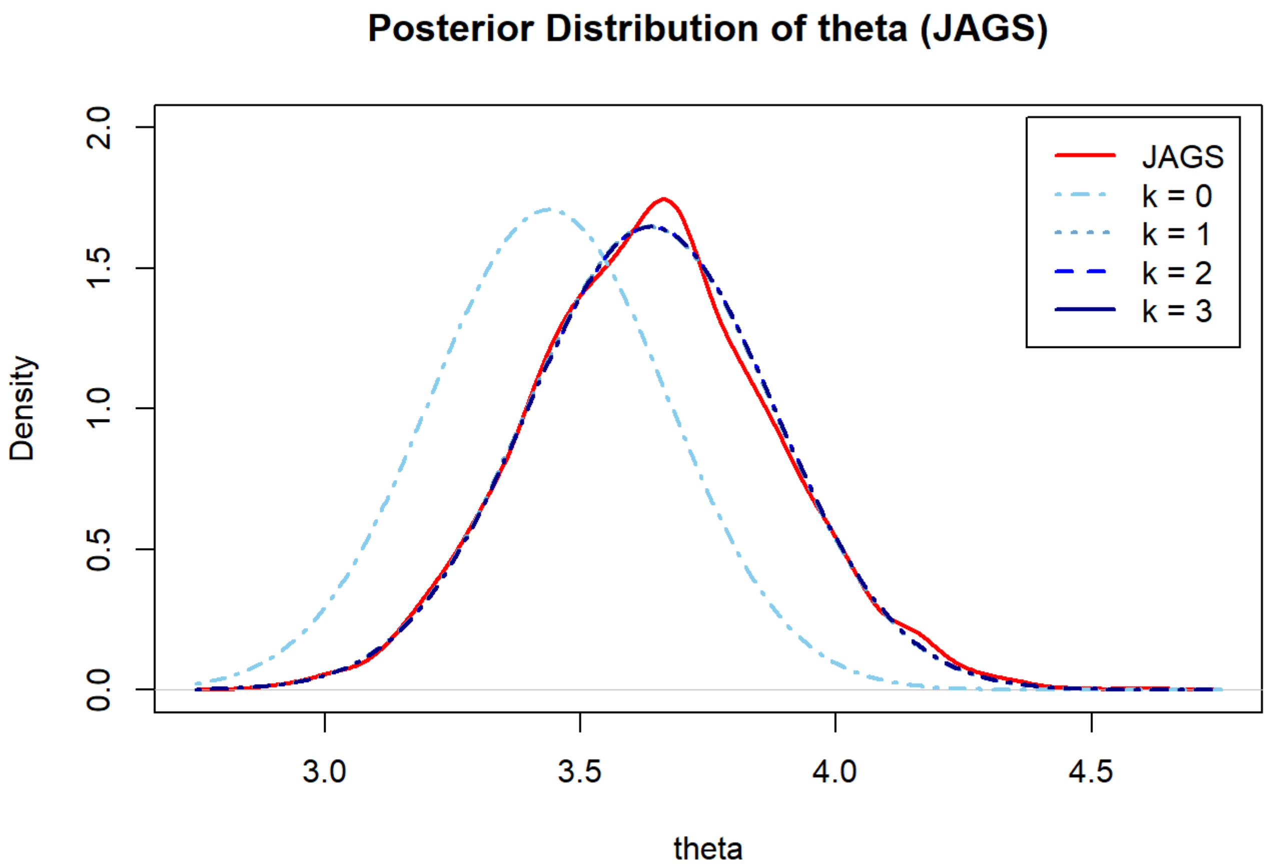





```
## Summary of theta samples
```

```
##
## Iterations = 10010:30000
## Thinning interval = 10
## Number of chains = 2
## Sample size per chain = 2000
##
## 1. Empirical mean and standard deviation for each variable,
##    plus standard error of the mean:
##
##           Mean             SD       Naive SE Time-series SE
##       3.637761       0.242047       0.003827       0.003827
##
## 2. Quantiles for each variable:
##
##  2.5%   25%   50%   75% 97.5%
## 3.175 3.474 3.636 3.795 4.139
```

- Compare the posterior distribution of θ with the results from part (d).



question - 3

Given the model $Y|n,p \sim \text{Binomial}(n,p)$ with prior distributions $n \sim \text{Poisson}(\lambda)$ and $p \sim \text{Beta}(a,b)$. The observed data is $Y = 10$.

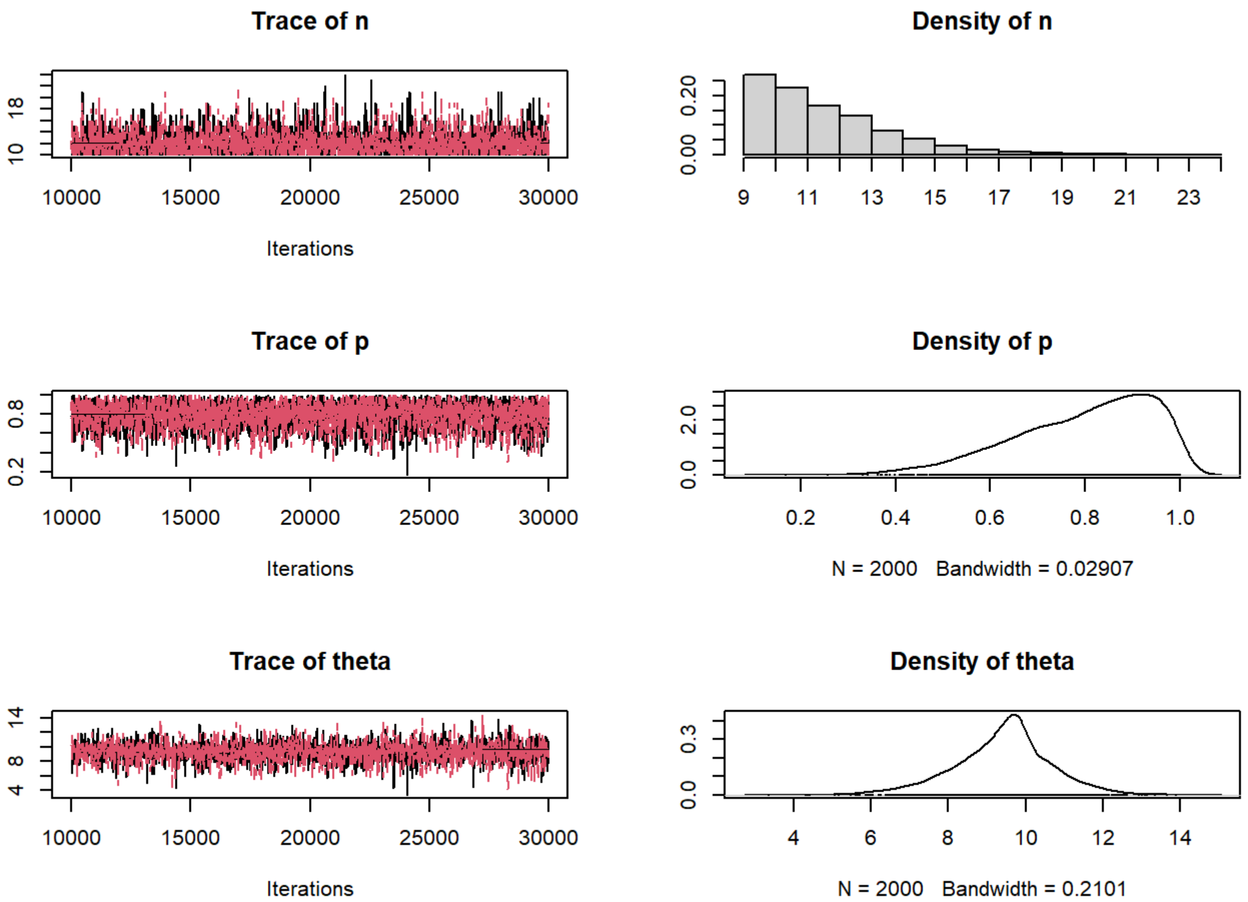
(a)

Convergence may be slow for this model due to several reasons:

- Correlation between parameters
- complex likelihood
- Inadequate Mixing
- Non-informative prior

(b)

Fit the model using JAGS with $\lambda = 10, a = b = 1$.

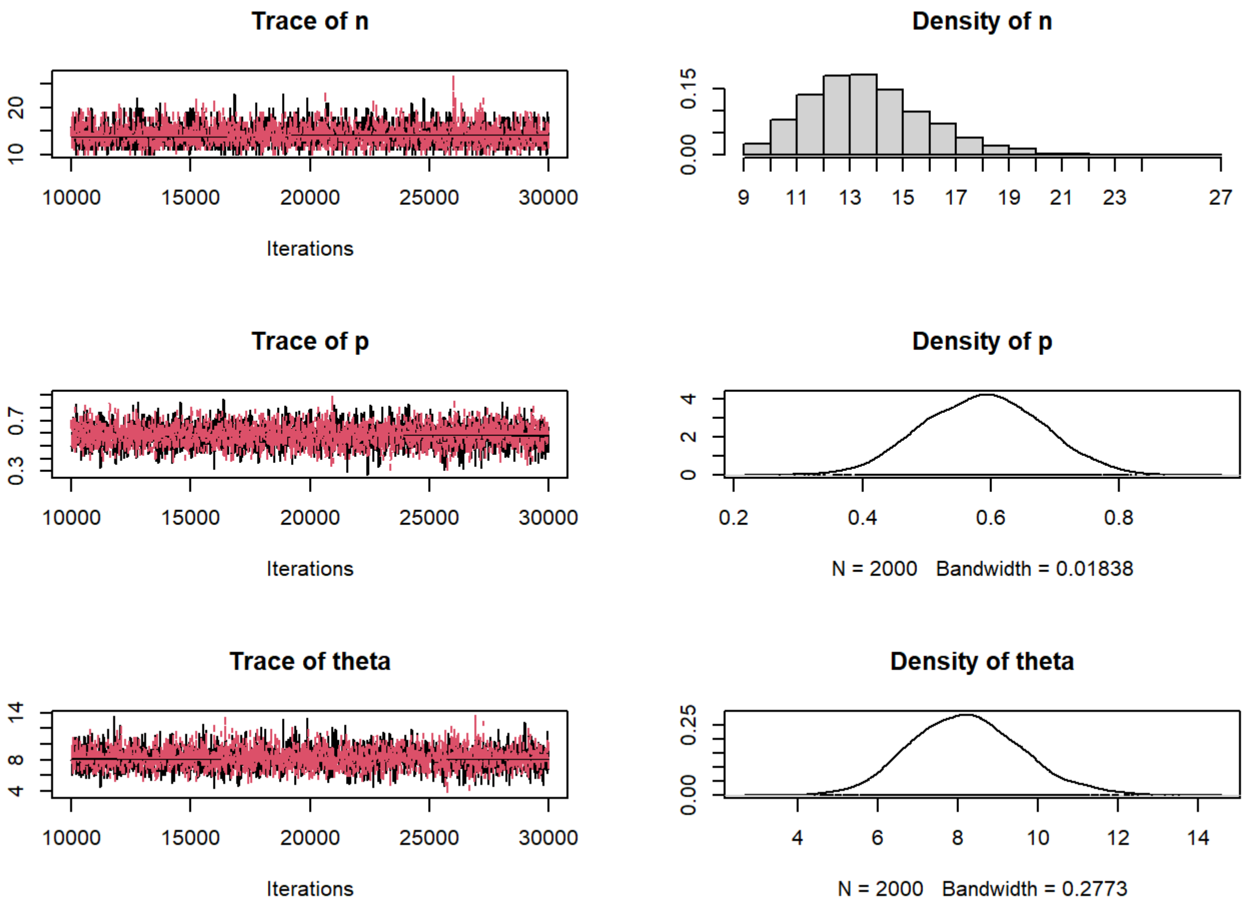



```
## Summary of samples when a, b = 1
```

```
##
## Iterations = 10010:30000
## Thinning interval = 10
## Number of chains = 2
## Sample size per chain = 2000
##
## 1. Empirical mean and standard deviation for each variable,
##    plus standard error of the mean:
##
##          Mean     SD Naive SE Time-series SE
## n     12.0370 2.0552 0.032495       0.031800
## p      0.7977 0.1441 0.002278       0.002278
## theta  9.3900 1.2682 0.020051       0.020053
##
## 2. Quantiles for each variable:
##
##          2.5%     25%     50%     75%   97.5%
## n     10.0000 10.0000 12.0000 13.0000 17.0000
## p      0.4701  0.7023  0.8241  0.9151  0.9915
## theta  6.5938  8.6845  9.5025 10.0799 11.8569
```

(c)

Fit the model using JAGS with $\lambda = 10, a = b = 10$.




```
## Summary of samples when a, b = 10
```

```
##
## Iterations = 10010:30000
## Thinning interval = 10
## Number of chains = 2
## Sample size per chain = 2000
##
## 1. Empirical mean and standard deviation for each variable,
##    plus standard error of the mean:
##
##          Mean      SD Naive SE Time-series SE
## n     14.1565 2.24138  0.03544        0.03412
## p      0.5869 0.09107  0.00144        0.00144
## theta  8.2219 1.37413  0.02173        0.02172
##
## 2. Quantiles for each variable:
##
##          2.5%     25%     50%   75%  97.5%
## n     10.0000 13.0000 14.0000 15.00 19.000
## p      0.4119  0.5232  0.5873  0.65  0.764
## theta  5.7039  7.2661  8.1707  9.11 11.121
```

Effect of Prior Distribution of p on Convergence

- For part (b), where $a = b = 1$, we use a non-informative prior for p. Non-informative priors generally have less influence on the posterior distribution, and convergence may be faster compared to informative priors.

- For part (c), where $a = b = 10$, we use a more informative prior for p. Informative priors can sometimes lead to slower convergence, especially if the prior distribution is significantly different from the likelihood. However, in this case, with a relatively large value of $a = b = 10$, the prior is still relatively diffuse and may not have a significant impact on convergence.
- Overall, the choice of prior distribution for p can affect convergence, especially if the prior distribution is very informative or if it conflicts with the likelihood. However, in this example, the effect of the prior distribution of p on convergence may be relatively minor due to the relatively non-informative nature of the priors used.

question - 4

Given a clinical trial gave six subjects a placebo and six subjects a new weight loss medication. The response variable is the change in weight (pounds) from baseline.

To Conduct a Bayesian analysis to compare the means of these two groups.

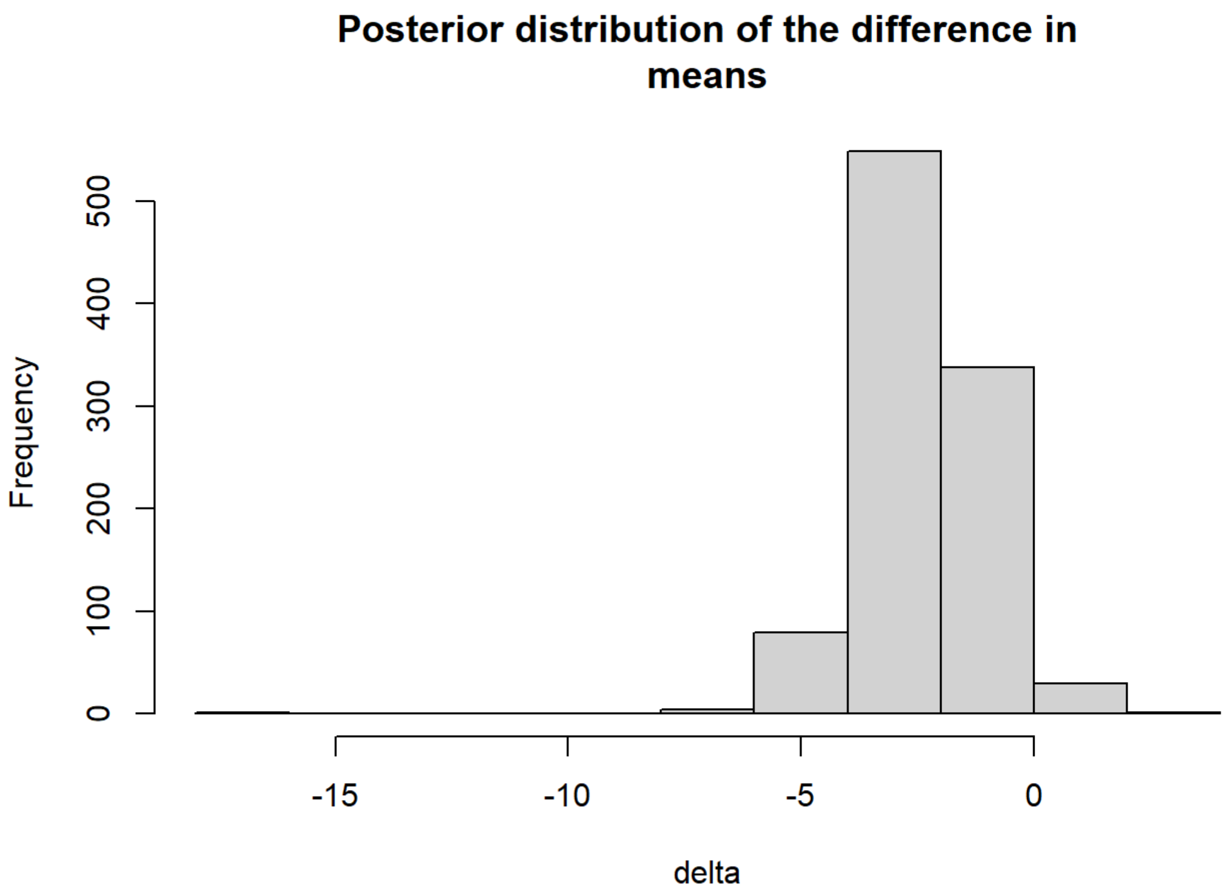
Let $Y_1 \sim \text{Normal}(\mu, \sigma^2)$ which the $n_1 = 6$ response of Placebo

Let $Y_2 \sim \text{Normal}(\mu + \delta, \sigma^2)$ which the $n_2 = 6$ response of Treatment.

The goal is to compare δ to 0.

```
## 95% credible interval when we assuming equal variance
```

```
## [1] -4.6058799 -0.1941201
```




```
## 95% credible interval when we assuming unequal variance
```

```
##       2.5%      97.5%
## -4.9406809  0.1330501
```

- The output of the code when we assume equal variance, the 95% credible interval is $[-4.605, -0.194]$.The output of the code when we assume unequal variance, the 95% credible interval is $[-4.940, 0.133]$.
- In equal variances case, we may say that treatment is effective as 0 is not included in its 95% credible interval, but in unequal variances case, we may say that treatment is not effective as 0 is included in its 95% credible interval.
- We can conclude that the interpretation of treatment effectiveness should be cautious because the intervals span values both above and below zero, indicating that the treatment may or may not be effective.
- It may be sensitive to the choice of prior, particularly if the prior distributions significantly influence the posterior distributions of the parameters.

question - 5

Given the dataset called **Boston**, in which the response variable (Y) is medv and the other 13 variables are covariates that describe the neighborhood.

$$
\begin{aligned}
Y_i &= \beta_0 + \sum_{j=1}^{p} X_{ij}\beta_j + \epsilon_i \\
Y &= X\beta + \epsilon
\end{aligned}
$$

$\text{No of observations} = i = 1, 2, \ldots\ldots, 506$
$\text{No of regression coefficients} = j = 1, 2, \ldots, p(= 13)$

(a)

Fit a Bayesian linear regression model with uninformative Gaussian priors for the regression coefficients.

$$
\begin{aligned}
\beta_0 &\sim \text{Normal}(0, 100^2) \\
\sigma^2 &\sim \text{IG}(0.01, 0.01) \\
\beta_1, \ldots. \beta_p &\sim \text{Normal}(0, 100^2) \text{ independent}
\end{aligned}
$$

```
## Loading required package: splines
```

```
## Loading required package: fds
```

```
## Loading required package: rainbow
```

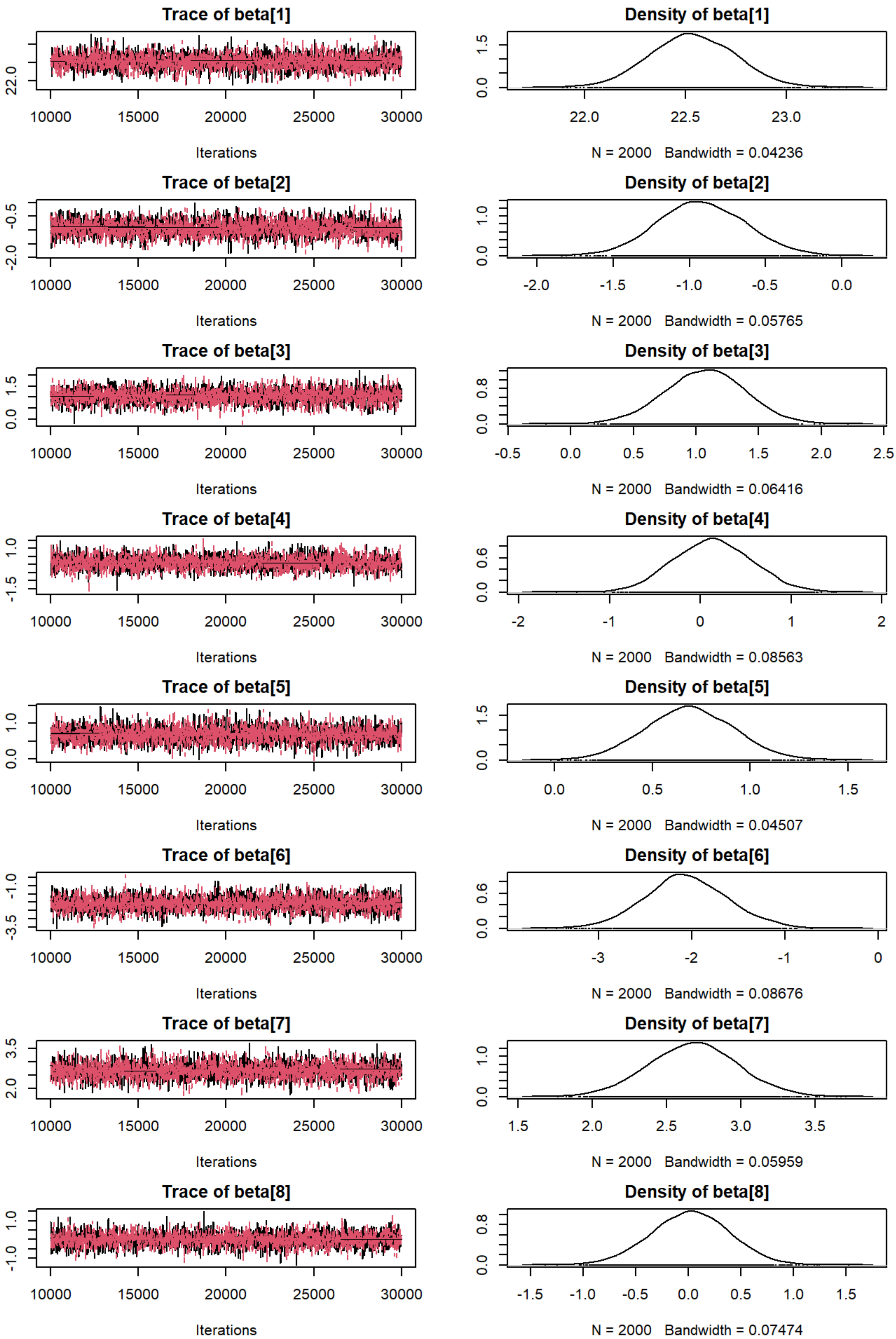
```
## Loading required package: pcaPP
```

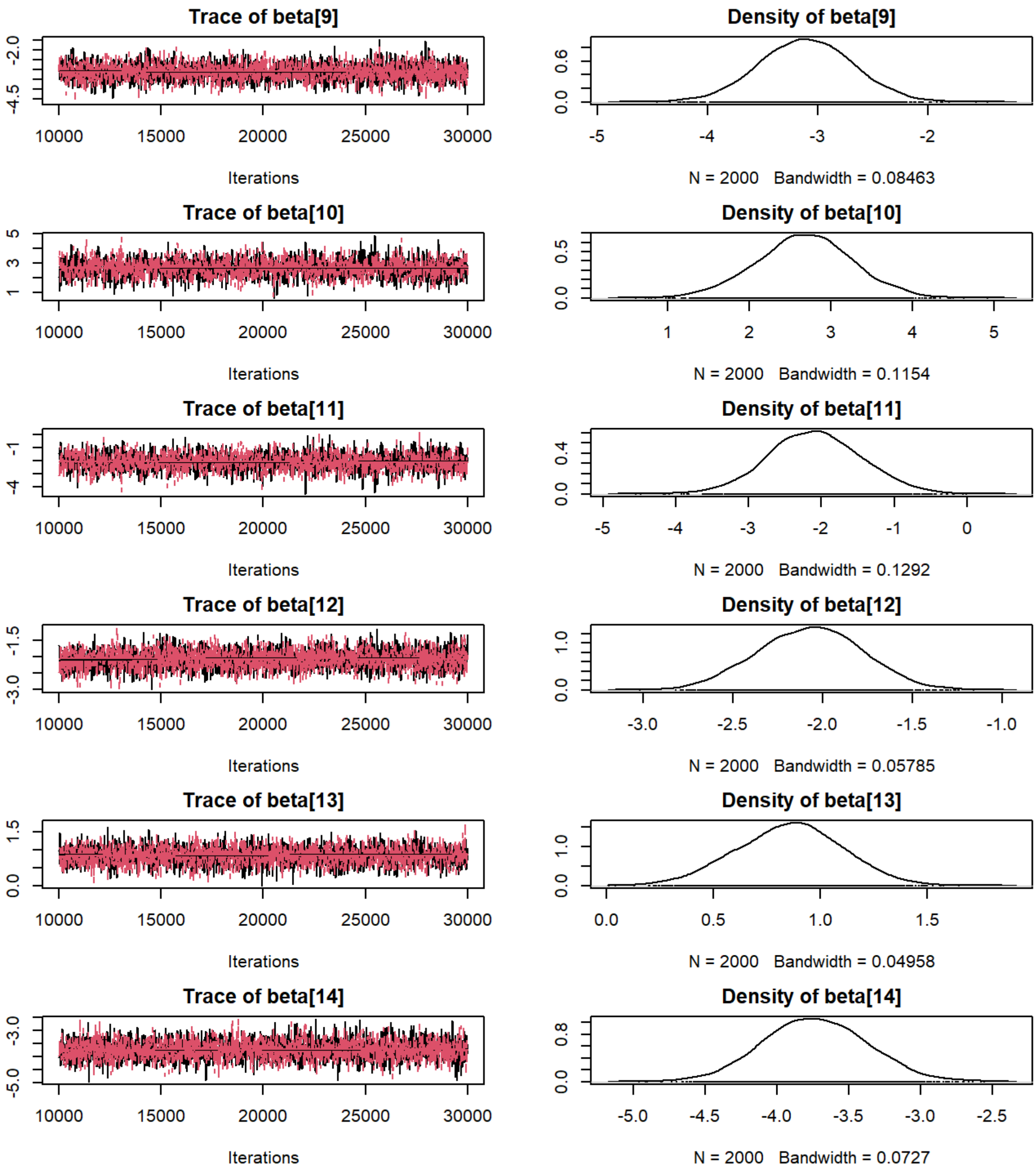
```
## Loading required package: RCurl
```

```
## Loading required package: deSolve
```

```
##
## Attaching package: 'fda'
```

```
## The following object is masked from 'package:graphics':
##
##     matplot
```


Trace of beta[1]
Iterations
Density of beta[1]
N = 2000 Bandwidth = 0.04236
Trace of beta[2]
Iterations
Density of beta[2]
N = 2000 Bandwidth = 0.05765
Trace of beta[3]
Iterations
Density of beta[3]
N = 2000 Bandwidth = 0.06416
Trace of beta[4]
Iterations
Density of beta[4]
N = 2000 Bandwidth = 0.08563
Trace of beta[5]
Iterations
Density of beta[5]
N = 2000 Bandwidth = 0.04507
Trace of beta[6]
Iterations
Density of beta[6]
N = 2000 Bandwidth = 0.08676
Trace of beta[7]
Iterations
Density of beta[7]
N = 2000 Bandwidth = 0.05959
Trace of beta[8]
Iterations
Density of beta[8]
N = 2000 Bandwidth = 0.07474




```
## Summary of beta samples with Gaussian priors
```

```
##
## Iterations = 10010:30000
## Thinning interval = 10
## Number of chains = 2
## Sample size per chain = 2000
##
## 1. Empirical mean and standard deviation for each variable,
##    plus standard error of the mean:
##
##              Mean     SD Naive SE Time-series SE
## beta[1]  22.53601 0.2099 0.003319       0.003319
## beta[2]  -0.92223 0.2857 0.004517       0.004394
## beta[3]   1.07008 0.3186 0.005038       0.004947
## beta[4]   0.13604 0.4243 0.006709       0.006846
## beta[5]   0.68979 0.2233 0.003531       0.003404
## beta[6]  -2.06283 0.4392 0.006944       0.006827
## beta[7]   2.67849 0.2953 0.004669       0.004261
## beta[8]   0.01664 0.3704 0.005856       0.005672
## beta[9]  -3.10906 0.4194 0.006631       0.006632
## beta[10]  2.67910 0.5860 0.009265       0.011410
## beta[11] -2.09000 0.6402 0.010122       0.012750
## beta[12] -2.07319 0.2867 0.004533       0.004533
## beta[13]  0.85394 0.2457 0.003885       0.003885
## beta[14] -3.73578 0.3603 0.005697       0.005346
##
## 2. Quantiles for each variable:
##
##             2.5%     25%      50%     75%   97.5%
## beta[1]  22.1280 22.3933 22.53226 22.6797 22.9518
## beta[2]  -1.4743 -1.1173 -0.92553 -0.7249 -0.3619
## beta[3]   0.4264  0.8598  1.07382  1.2859  1.6840
## beta[4]  -0.6720 -0.1577  0.13261  0.4271  0.9747
## beta[5]   0.2611  0.5395  0.68793  0.8425  1.1327
## beta[6]  -2.9102 -2.3479 -2.07237 -1.7718 -1.1647
## beta[7]   2.0948  2.4789  2.67949  2.8781  3.2607
## beta[8]  -0.7198 -0.2320  0.01855  0.2661  0.7227
## beta[9]  -3.9320 -3.3950 -3.11136 -2.8257 -2.2877
## beta[10]  1.5032  2.2982  2.68803  3.0647  3.8315
## beta[11] -3.3428 -2.5224 -2.09699 -1.6608 -0.8252
## beta[12] -2.6372 -2.2689 -2.06626 -1.8775 -1.5268
## beta[13]  0.3669  0.6893  0.86043  1.0218  1.3270
## beta[14] -4.4192 -3.9800 -3.74286 -3.4950 -3.0274
```

(b)

To Perform a classic least squares analysis.

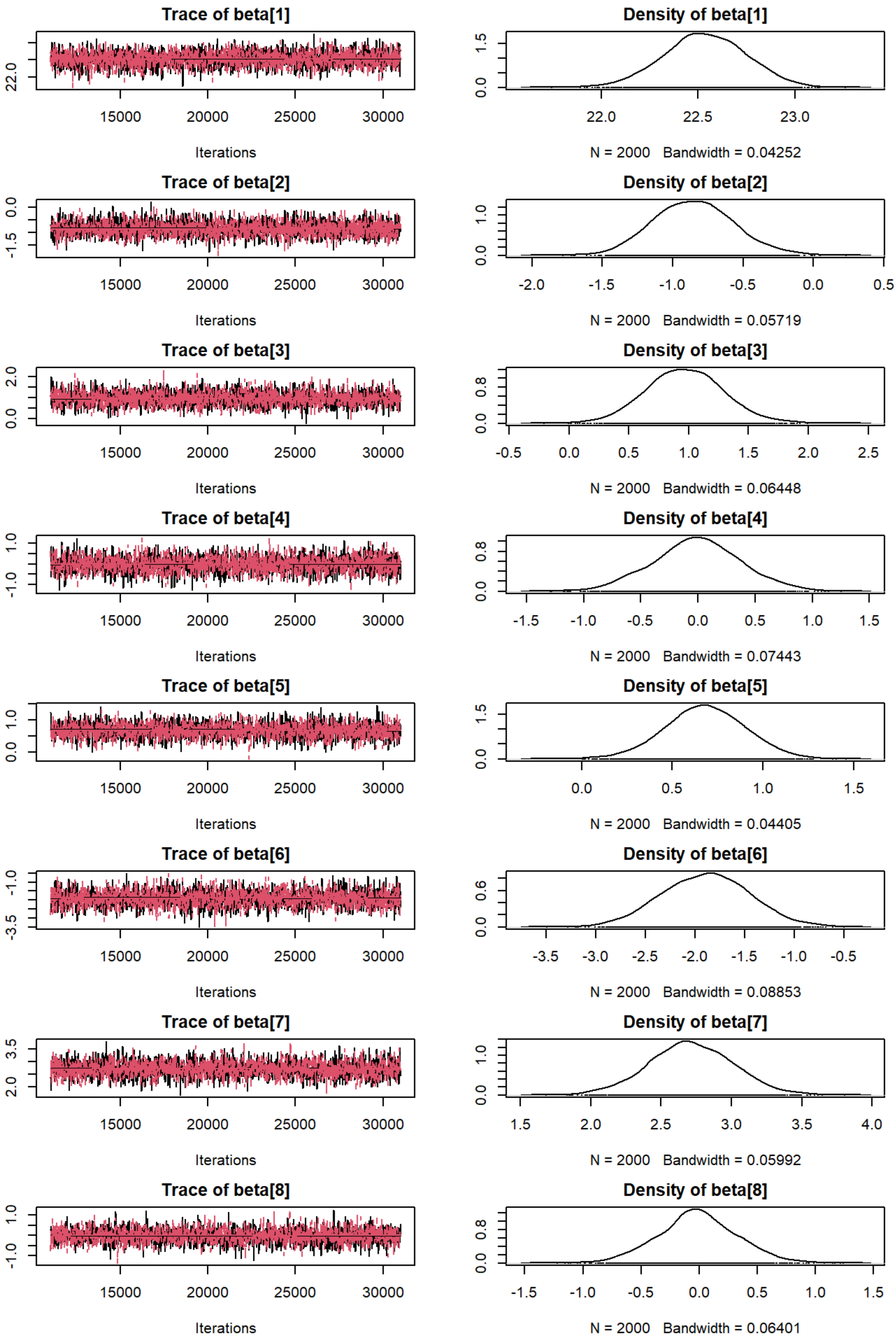
```
##
## Call:
## lm(formula = Y ~ x, data = Boston)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -15.595   -2.730   -0.518    1.777   26.199
##
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept) 22.53281    0.21095 106.814  < 2e-16 ***
## xcrim       -0.92906    0.28269  -3.287 0.001087 **
## xzn          1.08264    0.32016   3.382 0.000778 ***
## xindus       0.14104    0.42188   0.334 0.738288
## xchas        0.68241    0.21884   3.118 0.001925 **
## xnox        -2.05875    0.44262  -4.651 4.25e-06 ***
## xrm          2.67688    0.29364   9.116  < 2e-16 ***
## xage         0.01949    0.37184   0.052 0.958229
## xdis        -3.10712    0.41999  -7.398 6.01e-13 ***
## xrad         2.66485    0.57770   4.613 5.07e-06 ***
## xtax        -2.07884    0.63379  -3.280 0.001112 **
## xptratio    -2.06265    0.28323  -7.283 1.31e-12 ***
## xblack       0.85011    0.24521   3.467 0.000573 ***
## xlstat      -3.74733    0.36216 -10.347  < 2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 4.745 on 492 degrees of freedom
## Multiple R-squared:  0.7406, Adjusted R-squared:  0.7338
## F-statistic: 108.1 on 13 and 492 DF,  p-value: < 2.2e-16
```

- The values of regression coefficients in both Bayesian linear regression and classic least squares analysis are almost similar numerically and conceptually.

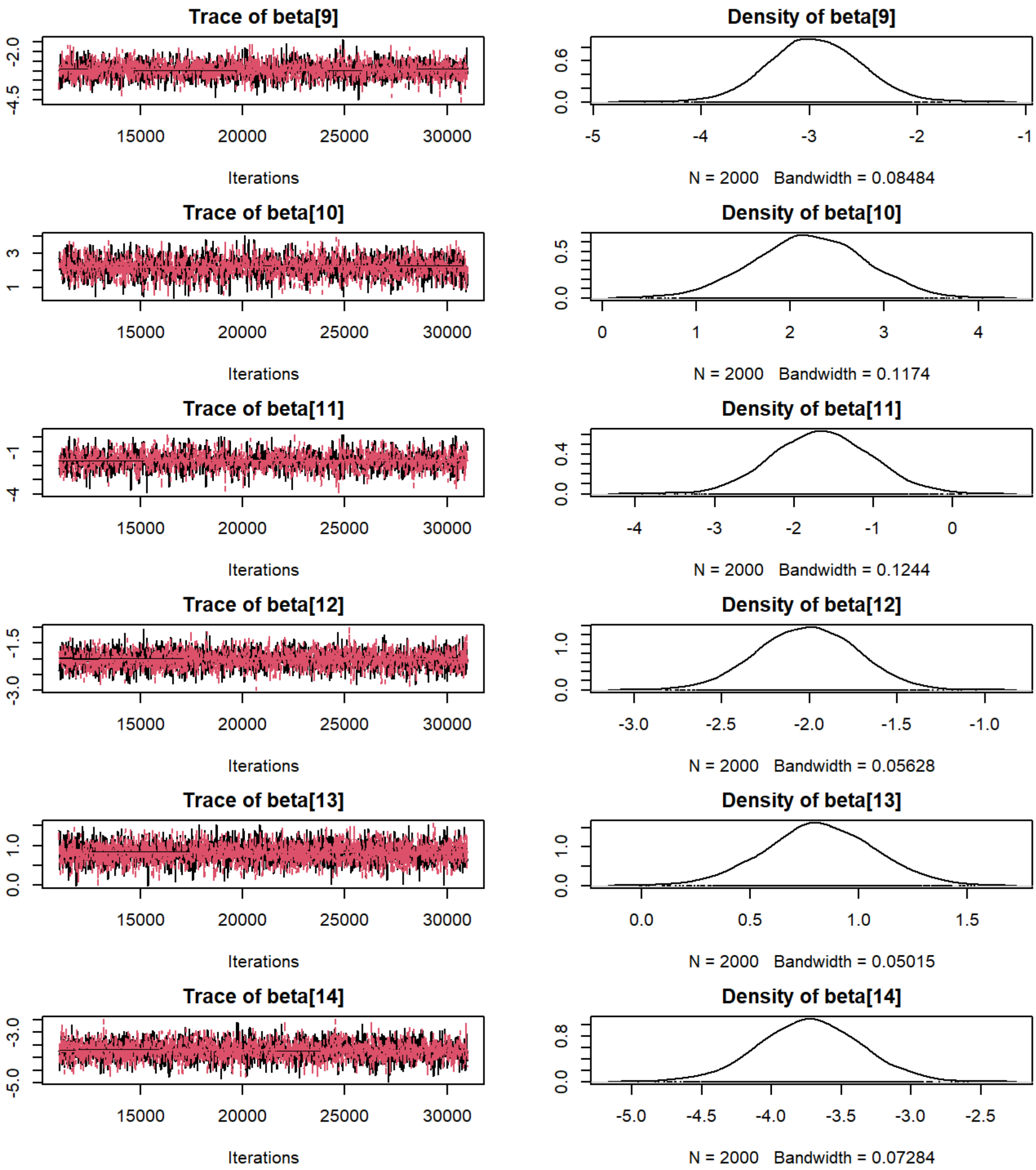
(c)

Fit the Bayesian model with double exponential priors for the regression coefficients.

$$\begin{aligned}
\beta_0 &\sim \text{Normal}(0, 100^2) \\
\sigma^2 &\sim \text{IG}(0.01, 0.01) \\
\beta_1, \ldots . \beta_p &\sim \text{Double Exponential}(0, \sigma^2 * \sigma_\beta^2) \text{ independent} \\
\sigma_\beta^2 &\sim \text{IG}(0.01, 0.01)
\end{aligned}$$



Trace of beta[1]
Iterations
Density of beta[1]
N = 2000 Bandwidth = 0.04252
Trace of beta[2]
Iterations
Density of beta[2]
N = 2000 Bandwidth = 0.05719
Trace of beta[3]
Iterations
Density of beta[3]
N = 2000 Bandwidth = 0.06448
Trace of beta[4]
Iterations
Density of beta[4]
N = 2000 Bandwidth = 0.07443
Trace of beta[5]
Iterations
Density of beta[5]
N = 2000 Bandwidth = 0.04405
Trace of beta[6]
Iterations
Density of beta[6]
N = 2000 Bandwidth = 0.08853
Trace of beta[7]
Iterations
Density of beta[7]
N = 2000 Bandwidth = 0.05992
Trace of beta[8]
Iterations
Density of beta[8]
N = 2000 Bandwidth = 0.06401




```
## Summary of beta samples with double exponential priors
```

```
##
## Iterations = 11010:31000
## Thinning interval = 10
## Number of chains = 2
## Sample size per chain = 2000
##
## 1. Empirical mean and standard deviation for each variable,
##    plus standard error of the mean:
##
##               Mean     SD Naive SE Time-series SE
## beta[1]  22.53306 0.2107 0.003332       0.003190
## beta[2]  -0.85214 0.2834 0.004482       0.004559
## beta[3]   0.96729 0.3217 0.005087       0.005166
## beta[4]  -0.01052 0.3819 0.006038       0.006626
## beta[5]   0.68008 0.2191 0.003465       0.003361
## beta[6]  -1.89107 0.4387 0.006937       0.007882
## beta[7]   2.71213 0.2969 0.004695       0.004880
## beta[8]  -0.01291 0.3393 0.005365       0.005602
## beta[9]  -2.95095 0.4220 0.006673       0.007506
## beta[10]  2.20105 0.5820 0.009203       0.014328
## beta[11] -1.65331 0.6165 0.009748       0.015278
## beta[12] -2.01367 0.2789 0.004410       0.004486
## beta[13]  0.82071 0.2485 0.003929       0.003775
## beta[14] -3.72290 0.3610 0.005708       0.005709
##
## 2. Quantiles for each variable:
##
##             2.5%     25%      50%     75%   97.5%
## beta[1]  22.1176 22.3937 22.53242 22.6790 22.9411
## beta[2]  -1.3778 -1.0510 -0.85620 -0.6638 -0.2706
## beta[3]   0.3490  0.7499  0.96060  1.1781  1.6102
## beta[4]  -0.7554 -0.2534 -0.01061  0.2408  0.7424
## beta[5]   0.2389  0.5358  0.67985  0.8284  1.1102
## beta[6]  -2.7603 -2.1911 -1.88448 -1.5895 -1.0343
## beta[7]   2.1277  2.5102  2.70640  2.9135  3.2973
## beta[8]  -0.6809 -0.2224 -0.01699  0.2026  0.6555
## beta[9]  -3.7643 -3.2292 -2.95893 -2.6658 -2.1268
## beta[10]  1.0497  1.8161  2.20220  2.6011  3.3128
## beta[11] -2.8476 -2.0787 -1.65272 -1.2389 -0.4408
## beta[12] -2.5620 -2.2039 -2.00989 -1.8194 -1.4769
## beta[13]  0.3262  0.6570  0.81906  0.9907  1.3061
## beta[14] -4.4468 -3.9653 -3.72409 -3.4797 -3.0213
```

Double Exponential Priors:

- The use of double exponential priors introduces regularization to the model, which can shrink coefficients towards zero and prevent overfitting.
- The summary of the fit with double exponential priors provides estimates of the regression coefficients and their uncertainties under the Bayesian framework.

Uninformative Gaussian Priors:

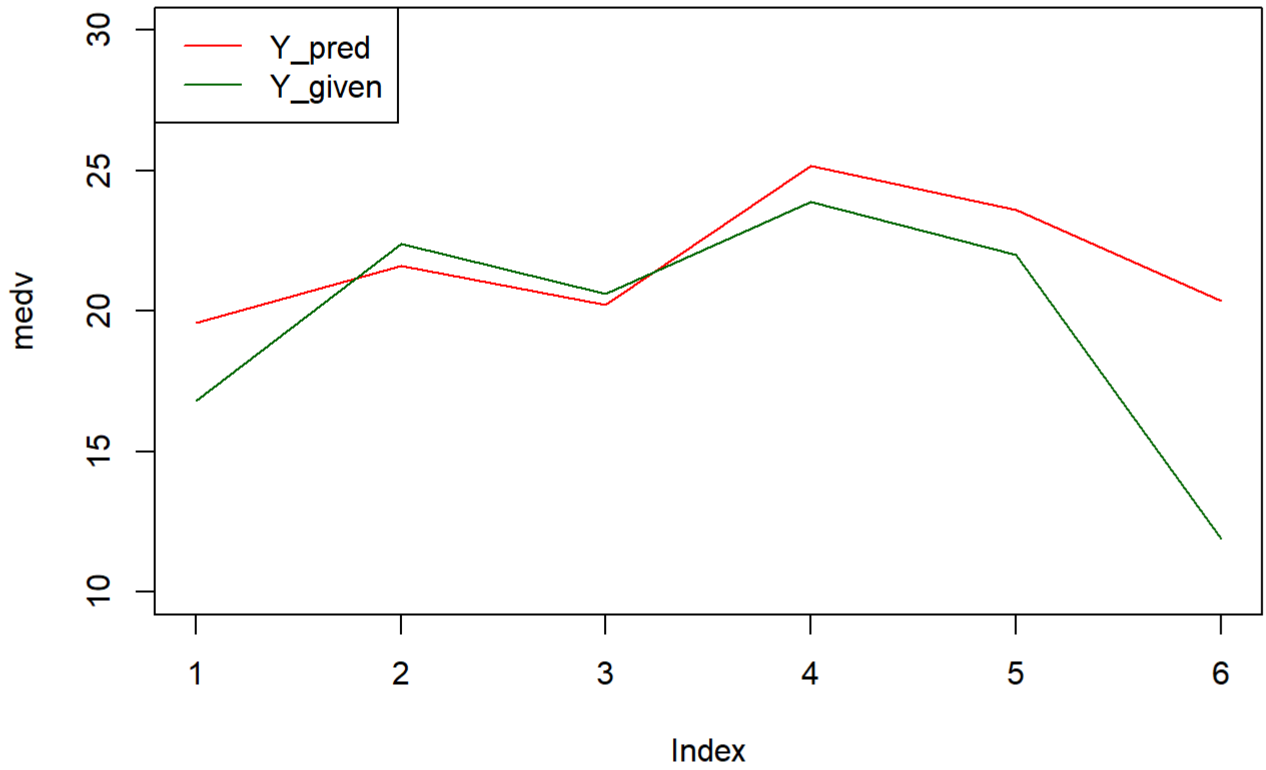
- Uninformative Gaussian priors provide minimal regularization and allow the data to dominate the inference process.

- The summary of the fit with uninformative Gaussian priors also provides estimates of the regression coefficients and their uncertainties.

(d)

Fit a Bayesian linear regression model in **(a)** using only the first 500 observations and find the posterior predictive distribution for the final 6 observations

```
##
## Iterations = 10010:30000
## Thinning interval = 10
## Number of chains = 2
## Sample size per chain = 2000
##
## 1. Empirical mean and standard deviation for each variable,
##    plus standard error of the mean:
##
##              Mean      SD  Naive SE Time-series SE
## beta[1]  22.29673 0.23464 3.710e-03      3.710e-03
## beta[2]  -0.96798 0.31088 4.915e-03      4.837e-03
## beta[3]   0.82957 0.35568 5.624e-03      5.624e-03
## beta[4]   0.10769 0.46363 7.331e-03      7.331e-03
## beta[5]   0.77707 0.24549 3.882e-03      3.882e-03
## beta[6]  -2.32774 0.48427 7.657e-03      7.393e-03
## beta[7]   2.64447 0.31620 5.000e-03      4.968e-03
## beta[8]  -0.09407 0.40630 6.424e-03      6.192e-03
## beta[9]  -2.66464 0.46471 7.348e-03      7.469e-03
## beta[10]  2.99455 0.62105 9.820e-03      1.238e-02
## beta[11] -1.69978 0.68373 1.081e-02      1.384e-02
## beta[12] -2.64986 0.31669 5.007e-03      4.934e-03
## beta[13]  0.84503 0.27184 4.298e-03      4.299e-03
## beta[14] -3.37172 0.39297 6.213e-03      5.903e-03
## taue      0.03528 0.00218 3.447e-05      3.531e-05
##
## 2. Quantiles for each variable:
##
##              2.5%      25%      50%     75%    97.5%
## beta[1]  21.83425 22.13953 22.29849 22.4529 22.77008
## beta[2]  -1.56770 -1.17771 -0.96898 -0.7565 -0.34417
## beta[3]   0.14632  0.59083  0.83016  1.0723  1.52556
## beta[4]  -0.82714 -0.20520  0.10500  0.4207  1.00611
## beta[5]   0.29951  0.61108  0.77959  0.9437  1.26742
## beta[6]  -3.27035 -2.66453 -2.32957 -1.9946 -1.37777
## beta[7]   2.02327  2.43302  2.64368  2.8582  3.26000
## beta[8]  -0.89487 -0.36646 -0.08989  0.1820  0.72327
## beta[9]  -3.56556 -2.98157 -2.66530 -2.3487 -1.75767
## beta[10]  1.78455  2.57150  2.98673  3.4036  4.22497
## beta[11] -3.07329 -2.14618 -1.69431 -1.2424 -0.39959
## beta[12] -3.27131 -2.86541 -2.65071 -2.4397 -2.03046
## beta[13]  0.31759  0.66181  0.83749  1.0289  1.37997
## beta[14] -4.13217 -3.63539 -3.37700 -3.1017 -2.61173
## taue      0.03106  0.03382  0.03528  0.0367  0.03965
```



- The predicated Y's are mostly same as the given Y's reasonably.