DATA7202 Assignment 3

Longpeng Xu

May 29, 2023

Contents

1 Question 1

(1) If $\theta \geq x_m$,

Prior: $\mathbb{P}(\theta) = \frac{\alpha x_m^\alpha}{\theta^{\alpha+1}}$

Given $\mathbb{P}(y_i|\theta) = \frac{1}{\theta}$, since Y_i's are iids,

$L(\theta) = \mathbb{P}(y|\theta) = \prod_{i=1}^{n} \mathbb{P}(y_i|\theta) = \frac{1}{\theta^n}$,

$\theta \geq \max\{x_m, Y_1, \cdots, Y_n\}$ since $L(\theta)$ is the joint of the uniforms

Posterior:

$$P(\theta|y) \propto \mathbb{P}(y|\theta)\mathbb{P}(\theta)$$

$$\propto \frac{1}{\theta^n} \frac{\alpha x_m^\alpha}{\theta^{\alpha+1}}$$

$\propto \frac{1}{\theta^{n+\alpha+1}}$, $\theta \geq \max\{x_m, Y_1, \dots, Y_n\}$, which is in the form of Pareto dist.

$\Rightarrow$ In this case, denote $\mathbb{P}(\theta|y) \sim \text{Pareto}(a, b) \Rightarrow f(\theta|a,b) = \frac{ab^a}{\theta^{a+1}}$, $\theta \geq b$

Compare $\frac{1}{\theta^{n+\alpha+1}}$ and $\frac{ab^a}{\theta^{a+1}} \propto \frac{1}{\theta^{a+1}}$ $\Rightarrow$ $a = n+\alpha$

Compare $\theta \geq \max\{x_m, Y_1, \cdots, Y_n\}$ and $\theta \geq b$ $\Rightarrow$ $b = \max\{x_m, Y_1, \cdots, Y_n\}$

(2) If $\theta < x_m$, $\mathbb{P}(\theta) = 0$, $\mathbb{P}(\theta|y) = \frac{\mathbb{P}(y|\theta)\mathbb{P}(\theta)}{\mathbb{P}(y)} = 0$, and naturally,

$\theta < x_m \leq \max\{x_m, Y_1, \cdots, Y_n\} \Rightarrow \theta < \max\{x_m, Y_1, \cdots, Y_n\}$

Overall, the posterior distribution of θ is $\mathbb{P}(\theta|y) \sim \text{Pareto}(n+\alpha, \max\{x_m, Y_1, \cdots, Y_n\})$

Hence, $p(\theta | n+\alpha, \max\{x_m, Y_1, \cdots, Y_m\})$

$$= \begin{cases} \frac{(n+\alpha)(\max\{x_m, Y_1, \cdots, Y_m\})^{n+\alpha}}{\theta^{n+\alpha+1}}, & \theta \geq \max\{x_m, Y_1, \cdots, Y_n\} \\ 0, & \theta < \max\{x_m, Y_1, \cdots, Y_n\} \end{cases}$$

2 Question 2

Prior $\mathbb{P}(p_1, \cdots, p_k) \propto \prod_{j=1}^{k} p_j^{\alpha_0^{(j)}-1}$

Given $\mathbb{P}(y_i | p_1, \cdots, p_k) = \prod_{j=1}^{k} p_j^{\mathbb{1}\{y_i=j\}}$, since Y_i's are iids,

$\mathbb{P}(y | p_1, \cdots, p_k) = \prod_{i=1}^{n} \prod_{j=1}^{k} p_j^{\mathbb{1}\{y_i=j\}}$

The posterior

$\mathbb{P}(p_1, \cdots, p_k | y)$

$\propto \mathbb{P}(y | p_1, \cdots, p_k) \cdot \mathbb{P}(p_1, \cdots, p_k)$

$\propto \prod_{i=1}^{n} \prod_{j=1}^{k} p_j^{\mathbb{1}\{y_i=j\}} \cdot \prod_{j=1}^{k} p_j^{\alpha_0^{(j)}-1}$

$\propto p_1^{\mathbb{1}\{Y_1=1\}} \cdots p_k^{\mathbb{1}\{Y_1=k\}} \cdots p_1^{\mathbb{1}\{Y_n=1\}} \cdots p_k^{\mathbb{1}\{Y_n=k\}} \cdot p_1^{\alpha_0^{(1)}-1} \cdots p_k^{\alpha_0^{(k)}-1}$

$\propto p_1^{(\mathbb{1}\{Y_1=1\} + \cdots \mathbb{1}\{Y_n=1\} + \alpha_0^{(1)} - 1)} \cdots p_k^{(\mathbb{1}\{Y_1=k\} + \cdots + \mathbb{1}\{Y_n=k\} + \alpha_0^{(k)} - 1)}$

$\propto \prod_{j=1}^{k} p_j^{(\sum_{i=1}^{n} \mathbb{1}\{Y_i=j\} + \alpha_0^{(j)} - 1)}$

which is the Dirichlet distribution with parameters $\alpha^{(j)} = \sum_{i=1}^{n} \mathbb{1}\{Y_i=j\} + \alpha_0^{(j)}$ for $j = 1, \cdots, k$.

Hence, the posterior distribution of $\theta = \{p_1, \cdots p_k\}$ is

$\mathbb{P}(p_1, \cdots p_k | y) \propto \text{Dirichlet}(\alpha^{(1)}, \cdots, \alpha^{(k)})$

where $\alpha^{(j)} = \sum_{i=1}^{n} \mathbb{1}\{Y_i=j\} + \alpha_0^{(j)}$ for $j = 1, \cdots, k$.

3 Question 3

(a)

$$f(X|Y=y) = \frac{f(x,y)}{f_Y(y)} = \frac{ce^{-(xy+x+y)}}{\int_0^\infty f(x,y)\,dx} \text{ where}$$

$$\int_0^\infty f(x,y)\,dx = \int_0^\infty ce^{-(xy+x+y)}\,dx$$

$$= ce^{-y}\int_0^\infty e^{-xy-x}\,dx$$

$$= ce^{-y}\int_0^\infty e^{-(y+1)x}\,dx$$

$$= ce^{-y}\cdot\frac{1}{-(y+1)}\cdot e^{-(y+1)x}\Big|_0^\infty$$

$$= \frac{ce^{-y}}{y+1} \Rightarrow f(X|Y=y) = \frac{ce^{-(xy+x+y)}}{\frac{ce^{-y}}{y+1}} = (y+1)e^{-x(y+1)}$$

which is the pdf of Expo$(y+1)$

$$f(Y|X=x) = \frac{f(x,y)}{f_X(x)} = \frac{ce^{-(xy+x+y)}}{\int_0^\infty f(x,y)\,dy} \text{ where}$$

$$\int_0^\infty f(x,y)\,dy = \int_0^\infty ce^{-(xy+x+y)}\,dy$$

$$= ce^{-x}\int_0^\infty e^{-xy-y}\,dy$$

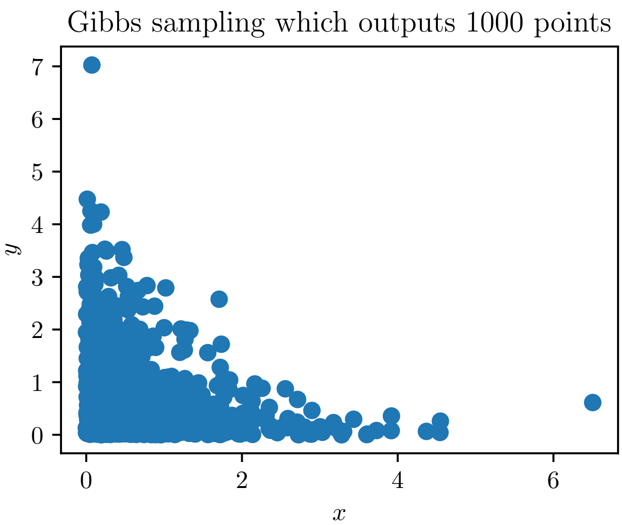
$$= ce^{-x}\int_0^\infty e^{-(x+1)y}\,dy$$

$$= ce^{-x}\cdot\frac{1}{-(x+1)}\cdot e^{-(x+1)y}\Big|_0^\infty$$

$$= \frac{ce^{-x}}{x+1} \Rightarrow f(Y|X=x) = \frac{ce^{-(xy+x+y)}}{\frac{ce^{-x}}{x+1}} = (x+1)e^{-y(x+1)}$$

which is the pdf of Expo$(x+1)$

(b) See the Python code for the Gibbs sampler in Appendix. The 1000 points from the output from the Gibbs sampler according to f:



4 Question 4

The `rjags, coda, MCMCvis` packages in R are used for this question.

(a) See the code in Appendix (R Code for Question 4 - Model 1: Q4 (a)).

(b) See the code in Appendix (R Code for Question 4 - Model 2: Q4 (b)).

(c) The results for Model 1: Page 5 to page 7. The results for Model 2: Page 8 to page 10.
See the code in Appendix (R Code for Question 4 - Model 1: Q4 (c) and Model 2: Q4(c))

Table 1 Information and diagnostics + Model Summary for Model 1

```
Information and diagnostics
===========================
Total iter:                       10000
Thin:                             1
Num chains:                       3
Max Rhat:                         1
Min n.eff:                        2121


Model summary
=============
              mean         sd        2.5%       50%      97.5% Rhat n.eff
alpha    3.4273524 1.69025179 1.09830426 3.1296014  7.5839097    1  2187
beta     5.2883759 2.66231039 1.61313620 4.8022707 11.8972312    1  2121
theta.1. 0.1229974 0.03313222 0.06542606 0.1205747  0.1945241    1 10406
theta.2. 0.2077938 0.02825973 0.15532961 0.2068358  0.2659247    1 16043
theta.3. 0.3025472 0.02632059 0.25246165 0.3019549  0.3551702    1 17808
theta.4. 0.4000575 0.02447907 0.35290189 0.3997859  0.4485610    1 17808
theta.5. 0.4969009 0.02854235 0.44104096 0.4967982  0.5533443    1 16725
theta.6. 0.5911161 0.03410109 0.52255192 0.5915354  0.6571018    1 17300
theta.7. 0.6759157 0.04662235 0.58180747 0.6774812  0.7641023    1 12348
```

Figure 1 The posterior distribution plot of Model 1

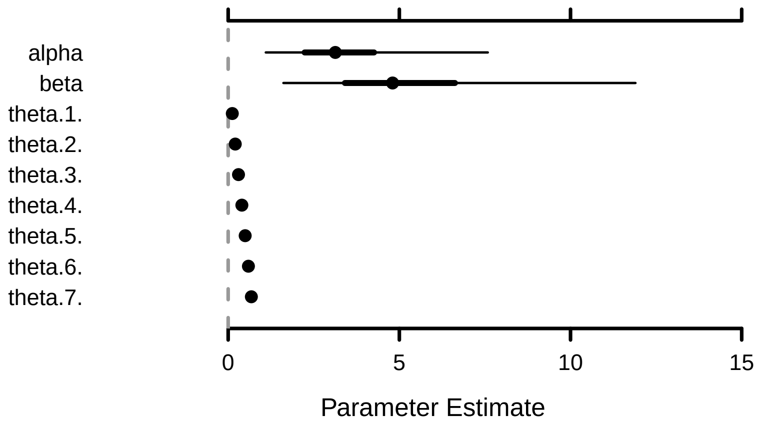
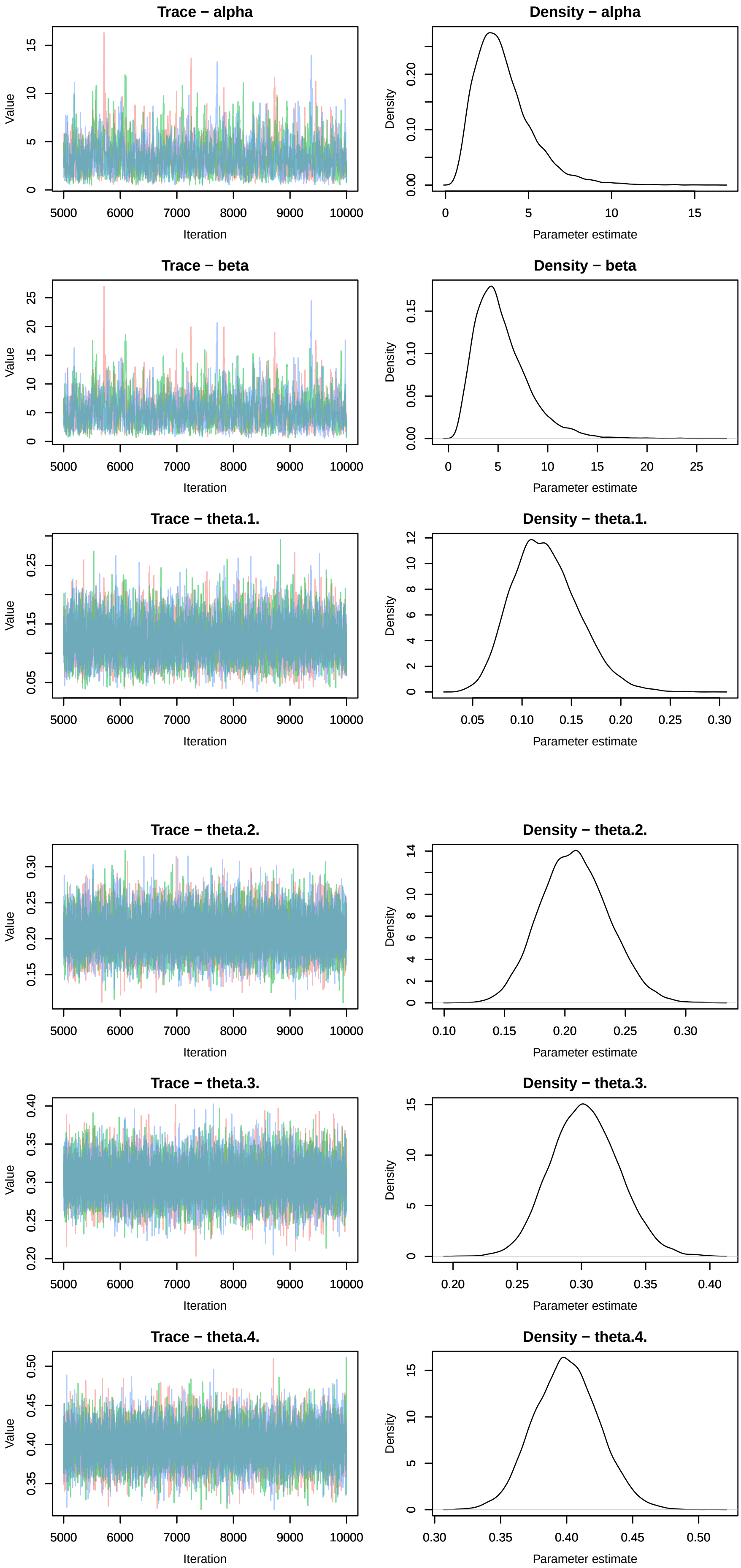


Figure 2 The trace plots of Model 1



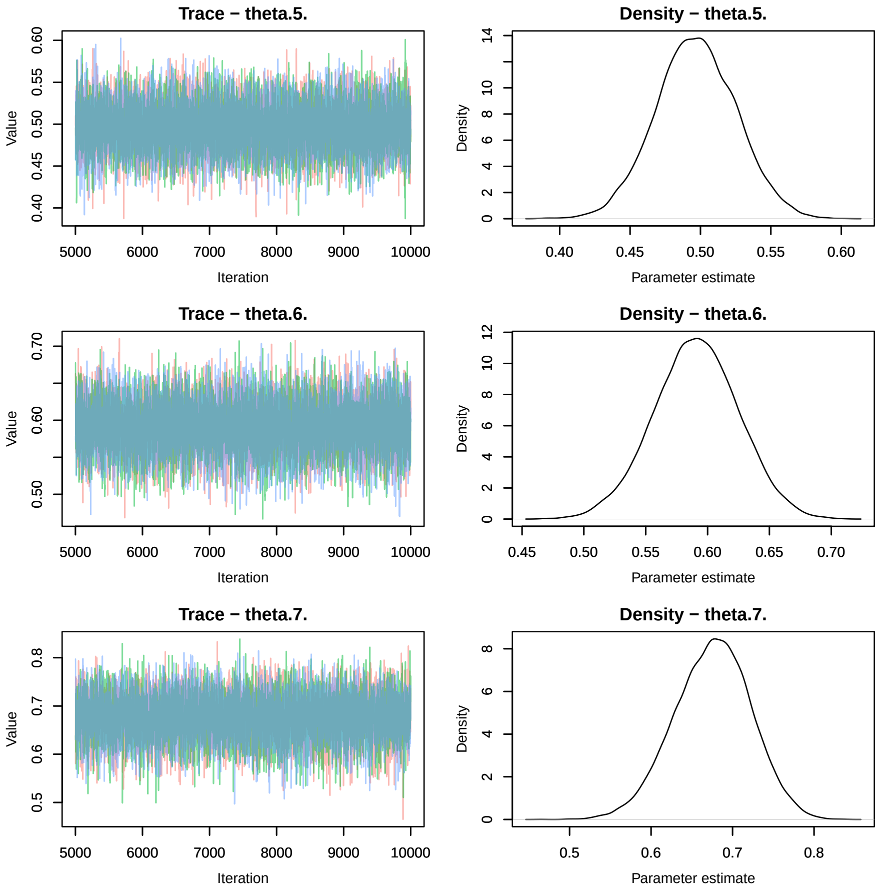

Trace − theta.5.
Density − theta.5.
Trace − theta.6.
Density − theta.6.
Trace − theta.7.
Density − theta.7.
Value
Iteration
Density
Parameter estimate
5000
6000
7000
8000
9000
10000

Table 2 Information and diagnostics + Model Summary for Model 2

```
Information and diagnostics
===========================
Total iter:                       10000
Thin:                             1
Num chains:                       3
Max Rhat:                         1.03
Min n.eff:                        873


Model summary
=============
                mean         sd          2.5%         50%       97.5% Rhat n.eff
beta0     1.39642497 0.23113200  0.9454121315  1.39129557  1.8787737 1.00   912
beta1    -0.46746382 0.05509435 -0.5861793696 -0.46472092 -0.3631526 1.00   873
sigma2    0.04767392 0.08379191  0.0006012456  0.02014141  0.2798794 1.03  1543
theta.1.  0.12692901 0.02343172  0.0792142145  0.12795941  0.1718794 1.00  1735
theta.2.  0.20003807 0.02171719  0.1579788742  0.19971689  0.2438503 1.00  4094
theta.3.  0.29232806 0.02154453  0.2515850342  0.29175673  0.3373011 1.00  7550
theta.4.  0.39478600 0.02048123  0.3561475201  0.39405481  0.4367149 1.00  8926
theta.5.  0.49986311 0.02379302  0.4524302988  0.49999422  0.5467612 1.00  7014
theta.6.  0.60671974 0.02824869  0.5497921850  0.60729422  0.6604232 1.00  3370
theta.7.  0.70998023 0.03374969  0.6398550709  0.71146187  0.7728927 1.00  2105
```

Figure 3 The posterior distribution plot of Model 1

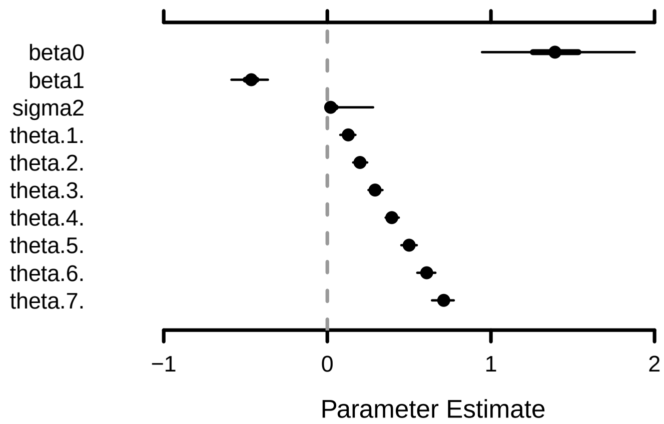
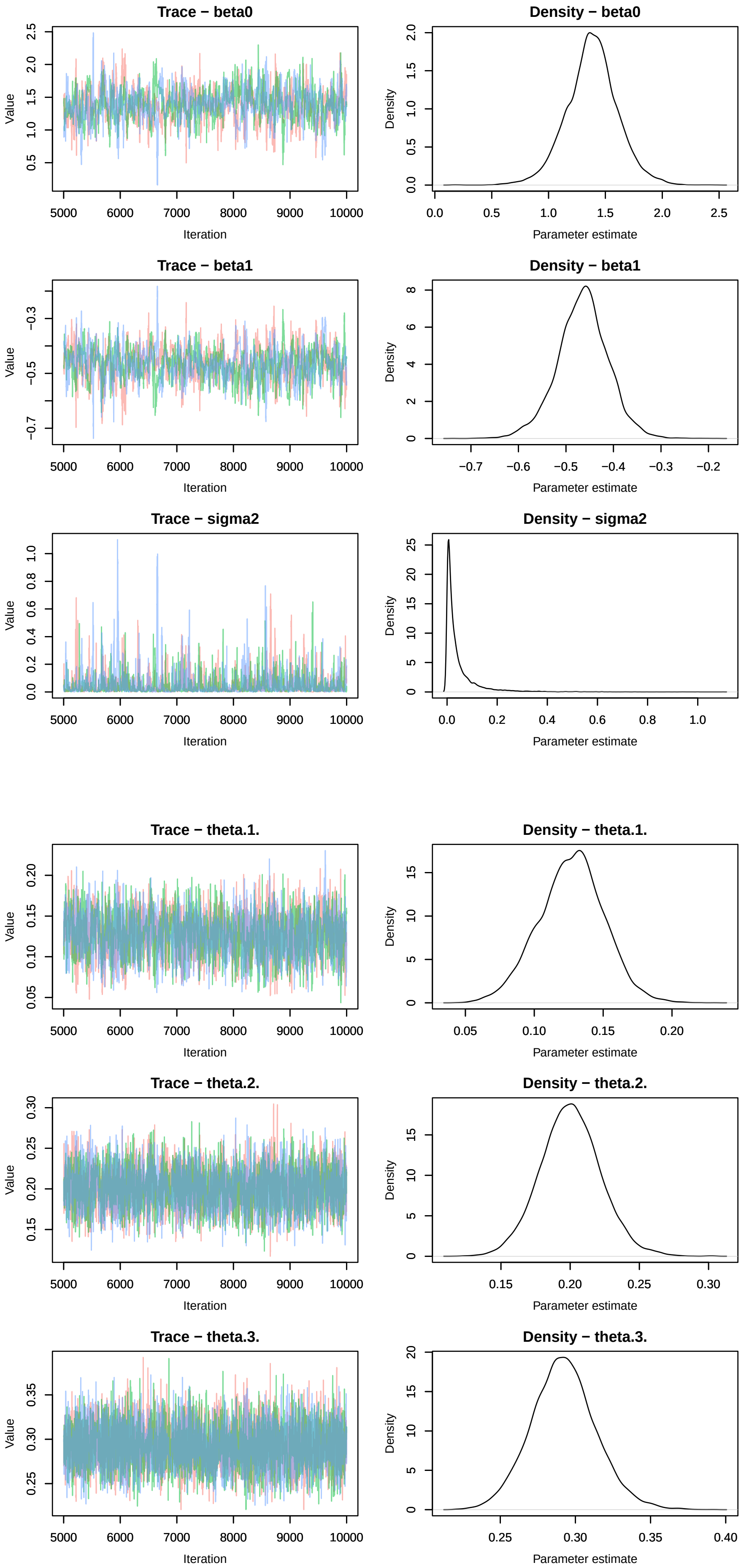
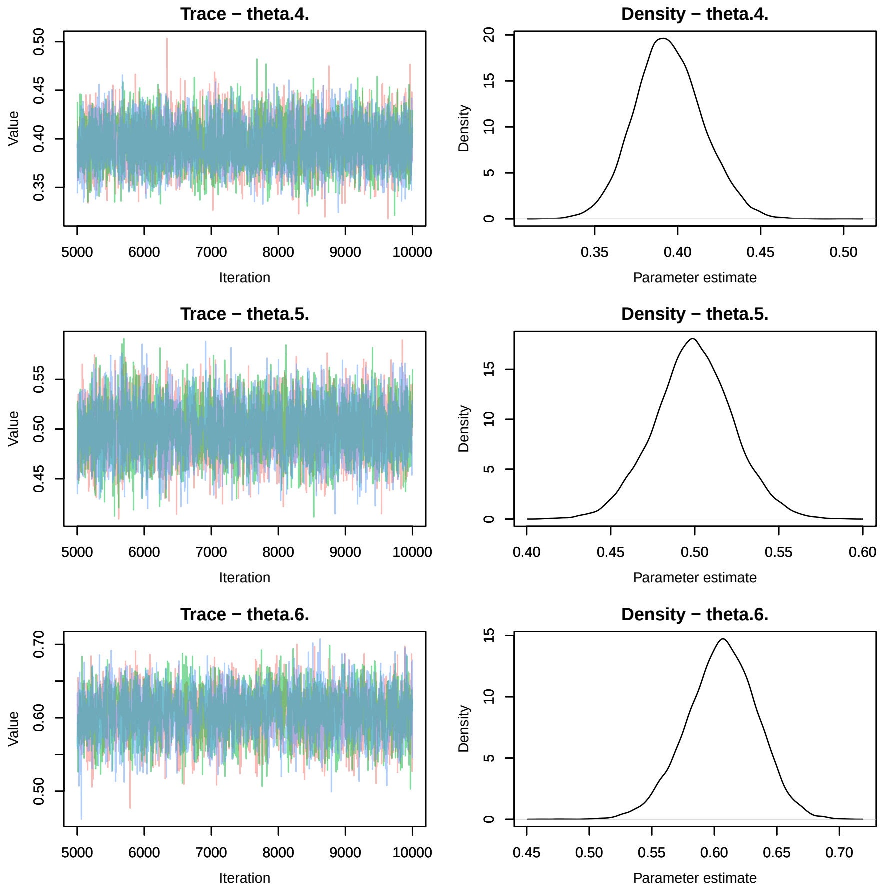
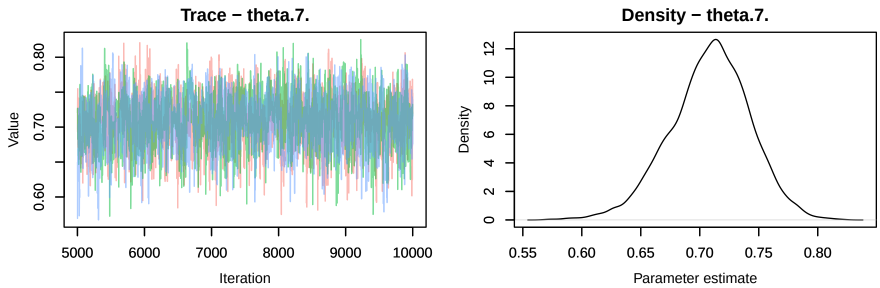


Figure 4 The trace plots of Model 2







(d) (1) β_1 adjusts the change in $\ln\left(\frac{\theta_j}{1-\theta_j}\right)$, the log odds of a YES reply, for a unit change in the distance.

(2) If $\beta_1 > 0$, the log odds of a YES reply increases (on average) as the distance increases.

(3) If $\beta_1 < 0$, the log odds of a YES reply decreases (on average) as the distance increases.

(4) If $\beta_1 > 0$ and β_1 increases, the log odds of a YES reply increases faster (on average) as the distance increases by the same amount as in (2).

(5) If $\beta_1 < 0$ and β_1 increases, the log odds of a YES reply decreases slower (on average) as the distance increases by the same amount as in (3).

Note 'on average' is specified for (2), (3), (4), (5) because the right hand side of

$$\ln\left(\frac{\theta_j}{1-\theta_j}\right) \sim N(\beta_0 + \beta_1 d_j, \sigma^2)$$

is not a deterministic function but a random one.

5 Appendix

Listing 1: **Python Code for Question 3**

```
# ==================================================================
# Question 3
# ==================================================================

import numpy as np

def gibbs(n=1000, burnin=100):
    # Start at an arbitrary point
    x, y = 1, 1
    samples_x = []
    samples_y = []

    for i in  range(n + burnin):
        # Update x given fixed y
        x = np.random.exponential(scale=1/(y+1))
        # Update y given fixed x
        y = np.random.exponential(scale=1/(x+1))

        # Collect samples after burn in
        if i >= burnin:
            samples_x += [x]
            samples_y += [y]

    return samples_x, samples_y


samples = gibbs()
samples

import matplotlib.pyplot as plt

plt.rcParams.update({'text.usetex': True, 'font.family': 'serif'})
plt.figure(figsize=(4,3),dpi=600, facecolor='white')
xpoints, ypoints = samples
plt.scatter(xpoints, ypoints)
plt.xlabel('$x$')
plt.ylabel('$y$')
plt.title('Gibbs sampling which outputs 1000 points')
plt.savefig('7202a3fig3.pdf')
plt.show()
```

Listing 2: **R Code for Question 4**

```r
# ============================================================================
# Question 4
# ============================================================================

setwd("D:/Current/DATA7202/a3")

library(coda)
library(rjags)

# ----------------------------------------------------------------------------
# Model 1: Q4(a)
# ----------------------------------------------------------------------------

# Obtain data for Model 1
data1 <-  list(
  J = 7,
  n = c(100, 200, 300, 400, 300, 200, 100),
  r = c(10, 40, 90, 160, 150, 120, 70))

# Create the sampler Model 1
model1 <- "
model {
    for (j in 1:J) {
        r[j] ~ dbin(theta[j], n[j])
        theta[j] ~ dbeta(alpha, beta)
    }
    alpha ~ dunif(0, 100)
    beta ~ dunif(0, 100)
}
"

# Prepare the sampler Model 1
jm1 <- rjags::jags.model(textConnection(model1), data = data1, n.chains = 3)

# Run the chains
jags_out1 <- rjags::coda.samples(jm1, variable.names = c('alpha','beta','theta'),
                                n. iter=10000, n.burnin=10000)

# Instantiate the MCMC object
samples1 <- coda::mcmc. list(jags_out1)
# Export to three csv files
write.csv(samples1[[1]], "D:/Current/DATA7202/a3/m1c1.csv", row.names = FALSE)
write.csv(samples1[[2]], "D:/Current/DATA7202/a3/m1c2.csv", row.names = FALSE)
write.csv(samples1[[3]], "D:/Current/DATA7202/a3/m1c3.csv", row.names = FALSE)

# ----------------------------------------------------------------------------
# Model 1: Q4(c)
# ----------------------------------------------------------------------------

library(MCMCvis)
```

```
m1c1 <- mcmc(read.csv("D:/Current/DATA7202/a3/m1c1.csv"))
m1c2 <- mcmc(read.csv("D:/Current/DATA7202/a3/m1c2.csv"))
m1c3 <- mcmc(read.csv("D:/Current/DATA7202/a3/m1c3.csv"))

mlist1 <- mcmc. list(m1c1, m1c2, m1c3)

MCMCdiag(mlist1)   # (info and diag) and model summary (.txt)
MCMCsummary(mlist1)# model summary
MCMCtrace(mlist1)  # (.pdf)
MCMCplot(mlist1)   # (export to .pdf)




# ----------------------------------------------------------------------------
# Model 2: Q4(b)
# ----------------------------------------------------------------------------

# Obtain data for Model 2
data2 <-  list(
  J = 7,
  n = c(100, 200, 300, 400, 300, 200, 100),
  d = c(7, 6, 5, 4, 3, 2, 1),
  r = c(10, 40, 90, 160, 150, 120, 70))

# Create the sampler Model 2
model2 <- "
model {
    for (j in 1:J) {
        r[j] ~ dbin(theta[j], n[j])
        logit(theta[j]) <- mu[j]
        mu[j] ~ dnorm(beta0 + beta1 * d[j], tau)
    }
    beta0 ~ dunif(-10, 10)
    beta1 ~ dunif(-10, 10)
    sigma2 ~ dunif(0, 100)
    tau <- pow(sigma2, -1)
}
"

# Prepare the sampler Model 2
jm2 <- rjags::jags.model(textConnection(model2), data = data2, n.chains = 3)

# Run the chains
jags_out2 <- rjags::coda.samples(jm2, variable.names = c('beta0','beta1','sigma2','theta'),
                                 n. iter=10000, n.burnin=10000)

# Instantiate the MCMC object
samples2 <- coda::mcmc. list(jags_out2)
# Export to three csv files
write.csv(samples2[[1]], "D:/Current/DATA7202/a3/m2c1.csv", row.names = FALSE)
```

```
write.csv(samples2[[2]], "D:/Current/DATA7202/a3/m2c2.csv", row.names = FALSE)
write.csv(samples2[[3]], "D:/Current/DATA7202/a3/m2c3.csv", row.names = FALSE)

# ------------------------------------------------------------------------------
# Model 2: Q4(c)
# ------------------------------------------------------------------------------

library(MCMCvis)

m2c1 <- mcmc(read.csv("D:/Current/DATA7202/a3/m2c1.csv"))
m2c2 <- mcmc(read.csv("D:/Current/DATA7202/a3/m2c2.csv"))
m2c3 <- mcmc(read.csv("D:/Current/DATA7202/a3/m2c3.csv"))

mlist2 <- mcmc. list(m2c1, m2c2, m2c3)

MCMCdiag(mlist2)   # (info and diag) and model summary (.txt)
MCMCsummary(mlist2)# model summary
MCMCtrace(mlist2)  # (.pdf)
MCMCplot(mlist2)   # (export to .pdf)
```