Assignment 1

JRGLAY001, GMDSBO006 and MGRKHE001

2025-04-11

Figure 1: Let me cook!

Question 1

a)

Model formulation

$$y = X\beta + e$$

where

$$e \sim \mathcal{N}(0, \sigma^2 I_n)$$

$$\mathbb{E}[y] = \mathbb{E}[X\beta + e] = \mathbb{E}[X\beta] + \mathbb{E}[e] = X\beta + 0 = X\beta$$
$$\mathrm{Var}[y] = \mathrm{Var}[X\beta + e] = \mathrm{Var}[X\beta] + \mathrm{Var}[e] + 2cov(X\beta, e) = 0 + \sigma^2 I_n + 0 = \sigma^2 I_n$$

Therefore

$$y \sim \mathcal{N}(X\beta, \sigma^2 I_n)$$

Density of Y

$$f_Y(y) = (2\pi\sigma^2)^{-\frac{n}{2}} exp(-\frac{1}{2\sigma^2}(y - X\beta)^T(y - X\beta))$$
$$f_Y(y) = (2\pi\sigma^2)^{-\frac{n}{2}} exp(-\frac{1}{2\sigma^2}(y^T y - 2\beta^T X^T y + \beta^T X^T X\beta))$$

Likelihood

$$L(\beta, y, X) \propto (\sigma^2)^{-\frac{n}{2}} exp(-\frac{1}{2\sigma^2}(y^T y - 2\beta^T X^T y + \beta^T X^T X\beta))$$

Since $M = M^{-1} = I_{k+1}$
Then the prior distribution is

$$[\beta|\sigma^2] \sim \mathcal{N}_{k+1}(\tilde{\beta}, \sigma^2 M^{-1})$$

Prior density

$$\pi(\beta|\sigma^2) = (2\pi)^{-\frac{k+1}{2}} \det(\sigma^2 M^{-1})^{-\frac{1}{2}} exp(-\frac{1}{2\sigma^2}(\beta - \tilde{\beta})^T M(\beta - \tilde{\beta}))$$
$$\pi(\beta|\sigma^2) = (2\pi)^{-\frac{k+1}{2}} \det(\sigma^2 M^{-1})^{-\frac{1}{2}} exp(-\frac{1}{2\sigma^2}(\beta^T M\beta - 2\beta^T M\tilde{\beta} + \tilde{\beta}^T M\tilde{\beta}))$$
$$\pi(\beta|\sigma^2) \propto exp(-\frac{1}{2\sigma^2}(\beta^T M\beta - 2\beta^T M\tilde{\beta} + \tilde{\beta}^T M\tilde{\beta}))$$

Prior distribution

$$[\sigma^2] \sim \mathcal{IG}(a, b)$$

Prior density

$$\pi(\sigma^2) = \frac{b^a}{\Gamma(a)} \sigma^{2(-a-1)} exp(-\frac{b}{\sigma^2})$$
$$\pi(\sigma^2) \propto \sigma^{2(-a-1)} exp(-\frac{b}{\sigma^2})$$

Assuming that σ^2 is known

Prior

$$\pi(\beta|\sigma^2) \propto exp(-\frac{1}{2\sigma^2}(\beta^T M\beta - 2\beta^T M\tilde{\beta}))$$

Likelihood

$$L(\beta|\sigma^2) \propto exp(-\frac{1}{2\sigma^2}(-2\beta^T X^T y + \beta^T X^T X\beta))$$

Posterior

$$\pi(\beta|y,X) \propto exp(-\frac{1}{2\sigma^2}(\beta^T(M + X^T X)\beta - 2\beta^T(X^T y + M\tilde{\beta})))$$

Recall that $(X^T X)\hat{\beta} = X^T y$

$$\pi(\beta|y,X) \propto exp(-\frac{1}{2\sigma^2}(\beta^T(M + X^T X)\beta - 2\beta^T((X^T X)\hat{\beta} + M\tilde{\beta})))$$

Let $A = M + X^T X$ and $b = (X^T X)\hat{\beta} + M\tilde{\beta}$

Completing the square using the following identity

$$\beta^T A\beta - 2\beta^T b = (\beta - \mu_\beta)^T A(\beta - \mu_\beta) - \mu_\beta^T A\mu_\beta$$

where $\mu_\beta = A^{-1}b$

$$\pi(\beta|y,X) \propto exp(-\frac{1}{2\sigma^2}((\beta - \mu_\beta)^T A(\beta - \mu_\beta) - \mu_\beta^T A\mu_\beta))$$

but $\mu_\beta^T A\mu_\beta$ is independent of β

$$\pi(\beta|y,X) \propto exp(-\frac{1}{2\sigma^2}(\beta - \mu_\beta)^T A(\beta - \mu_\beta))$$

$$\pi(\beta|y,X) \propto exp(-\frac{1}{2\sigma^2}(\beta - \mu_\beta)^T(M + X^T X)(\beta - \mu_\beta))$$

meaning

$$[\beta|\sigma^2, y, X] \sim \mathcal{N}_{k+1}(\mu_\beta, \sigma^2(M + X^T X)^{-1})$$

where

$$\mu_\beta = (M + X^T X)^{-1}((X^T X)\hat{\beta} + M\tilde{\beta})$$

Joint distribution

$$J = [\beta, \sigma^2 | y, X]$$

Consider

$$P = L(\beta, y, X)[\beta|\sigma^2]$$

$$P \propto (\sigma^2)^{-\frac{n}{2}} exp(-\frac{1}{2\sigma^2}(y^T y - 2\beta^T X^T y + \beta^T X^T X \beta))(\sigma^2)^{-\frac{k+1}{2}} exp(-\frac{1}{2\sigma^2}(\beta^T M \beta - 2\beta^T M \tilde{\beta} + \tilde{\beta}^T M \tilde{\beta}))$$

$$P \propto (\sigma^2)^{-\frac{n+k+1}{2}} exp(-\frac{1}{2\sigma^2}(y^T y + \tilde{\beta}^T M \tilde{\beta})) exp(-\frac{1}{2\sigma^2}(\beta^T (M + X^T X)\beta - 2\beta^T((X^T X)\hat{\beta} + M\tilde{\beta}))$$

Let $A = M + X^T X$ and $b = (X^T X)\hat{\beta} + M\tilde{\beta}$

Completing the square using the following identity

$$\beta^T A \beta - 2\beta^T b = (\beta - \mu_\beta)^T A (\beta - \mu_\beta) - \mu_\beta^T A \mu_\beta$$

where $\mu_\beta = A^{-1} b$

$$P \propto (\sigma^2)^{-\frac{n+k+1}{2}} exp(-\frac{1}{2\sigma^2}(y^T y + \tilde{\beta}^T M \tilde{\beta})) exp(-\frac{1}{2\sigma^2}((\beta - \mu_\beta)^T A (\beta - \mu_\beta) - \mu_\beta^T A \mu_\beta))$$

$$P \propto (\sigma^2)^{-\frac{n+k+1}{2}} exp(-\frac{1}{2\sigma^2}(y^T y + \tilde{\beta}^T M \tilde{\beta} - \mu_\beta^T A \mu_\beta)) exp(-\frac{1}{2\sigma^2}(\beta - \mu_\beta)^T A (\beta - \mu_\beta))$$

$$[\sigma^2] \propto \sigma^{2(-a-1)} exp(-\frac{b}{\sigma^2})$$

$$J \propto L(\beta, y, X)[\beta|\sigma^2][\sigma^2]$$

$$J \propto (\sigma^2)^{-(\frac{n+k+1}{2}+a+1)} exp(-\frac{1}{2\sigma^2}(y^T y + \tilde{\beta}^T M \tilde{\beta} - \mu_\beta^T A \mu_\beta)) exp(-\frac{1}{2\sigma^2}(\beta - \mu_\beta)^T A (\beta - \mu_\beta)) exp(-\frac{b}{\sigma^2})$$

Let $A_2 = y^T y + \tilde{\beta}^T M \tilde{\beta} - \mu_\beta^T A \mu_\beta$

$$J \propto (\sigma^2)^{-(\frac{n+k+1}{2}+a+1)} exp(-\frac{1}{\sigma^2}(b + \frac{A_2}{2})) exp(-\frac{1}{2\sigma^2}(\beta - \mu_\beta)^T A (\beta - \mu_\beta))$$

$$[\sigma^2 | y, X] = \int_\beta [\beta, \sigma^2 | y, X] \, d\beta$$

$$[\sigma^2 | y, X] \propto (\sigma^2)^{-(\frac{n+k+1}{2}+a+1)} exp(-\frac{1}{\sigma^2}(b + \frac{A_2}{2})) \int_\beta exp(-\frac{1}{2\sigma^2}(\beta - \mu_\beta)^T A (\beta - \mu_\beta) \, d\beta$$

Recall

$$\int_\beta (2\pi)^{-\frac{k+1}{2}} det((\frac{1}{\sigma^2}(M + X^T X))^{-1})^{-\frac{1}{2}} exp(-\frac{1}{2\sigma^2}(\beta - \mu_\beta)^T A (\beta - \mu_\beta)) \, d\beta = 1$$

then

$$I = \int_\beta exp(-\frac{1}{2\sigma^2}(\beta - \mu_\beta)^T A (\beta - \mu_\beta) \, d\beta = (2\pi)^{\frac{k+1}{2}} det((\frac{1}{\sigma^2}(M + X^T X))^{-1})^{\frac{1}{2}})$$

We know that $det(aA) = a^k det(A)$, where A is a k by k matrix.

$$I = (2\pi)^{\frac{k+1}{2}} (\sigma^2)^{\frac{k+1}{2}} det((M + X^T X)^{-1})^{\frac{1}{2}}$$

$$I \propto (\sigma^2)^{\frac{k+1}{2}}$$

$$[\sigma^2 | y, X] \propto (\sigma^2)^{-(\frac{n+k+1}{2}+a+1)} exp(-\frac{1}{\sigma^2}(b + \frac{A_2}{2}))(\sigma^2)^{\frac{k+1}{2}}$$

$$[\sigma^2 | y, X] \propto (\sigma^2)^{-(\frac{n}{2}+a)-1} exp(-\frac{1}{\sigma^2}(b + \frac{A_2}{2}))$$

meaning

$$[\sigma^2 | y, X] \sim \mathcal{IG}(a + \frac{n}{2}, b + \frac{A_2}{2})$$

where $a = 1$, $b = 1$ and $A_2 = y^T y + \tilde{\beta}^T M \tilde{\beta} - \mu_\beta^T (M + X^T X) \mu_\beta$

b)

Table 1: First six rows of σ^2 sample values

σ^2
1.968357
1.695290
1.674221
1.982145
1.534436
1.750102

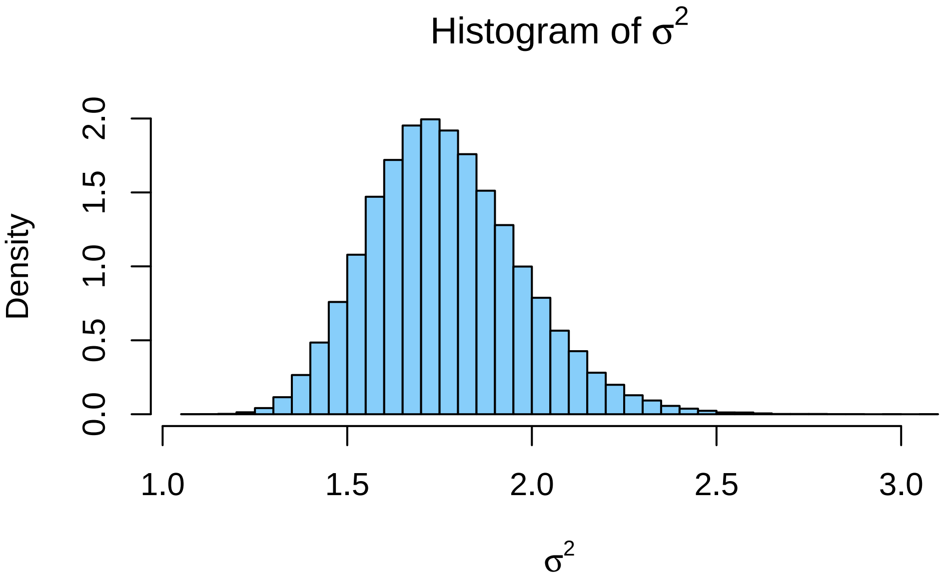
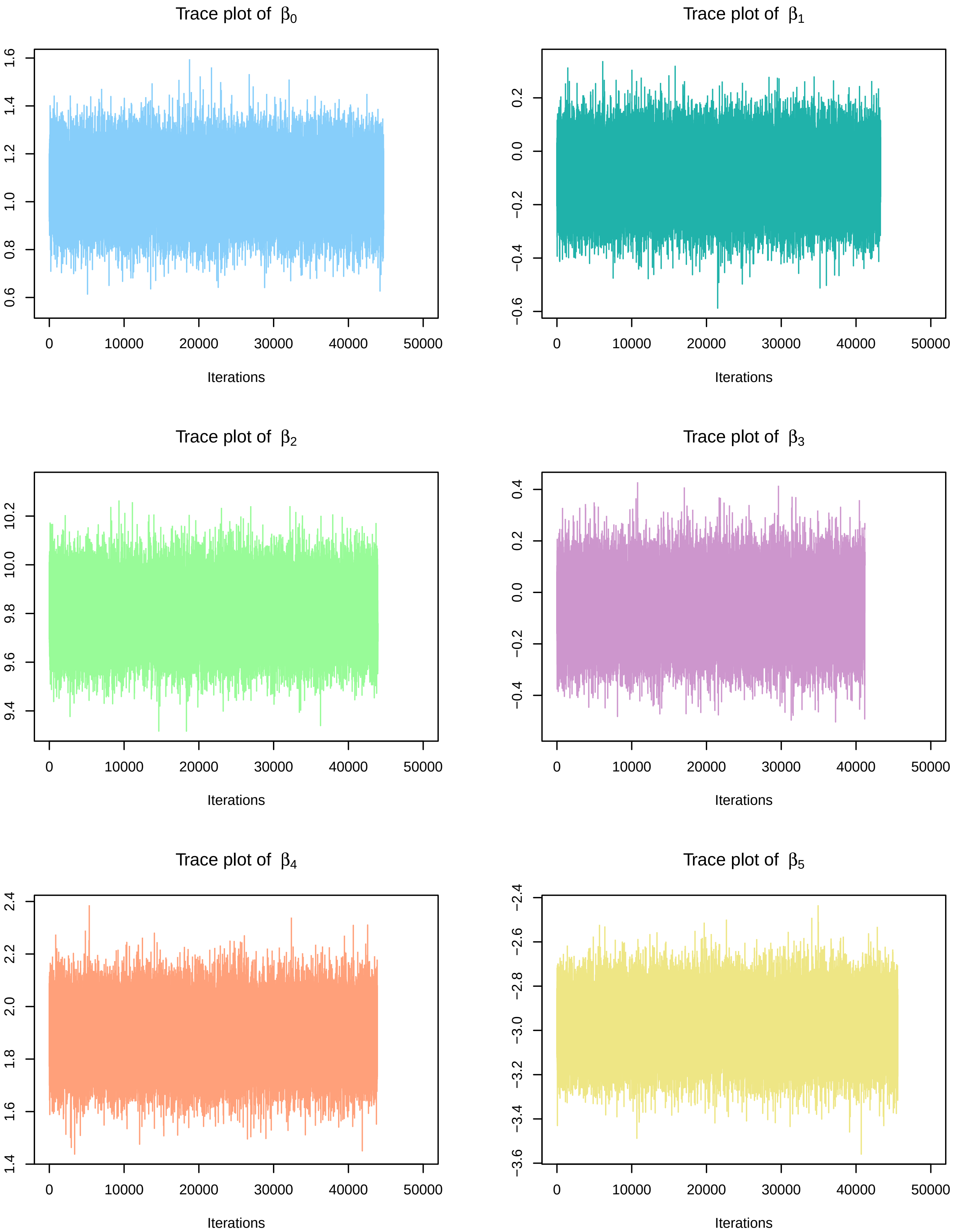


Table 2: First six rows of the β sample values

β_0	β_1	β_2	β_3	β_4	β_5
0.9227484	-0.10623460	9.759792	0.10381837	2.051554	-2.839680
1.0567194	-0.20497249	9.788940	-0.09386065	1.772719	-2.809105
1.0122741	-0.04329813	9.801656	0.05378141	2.089081	-2.834065
1.2013402	-0.15871660	9.956266	-0.15968348	1.990569	-2.957680
1.1146270	-0.13473328	9.724995	0.06301181	1.879163	-3.124324
0.9180509	-0.14602992	10.011397	0.08100581	1.804324	-2.854732
1.0698329	0.01746358	9.703786	-0.08617031	1.880744	-2.950838

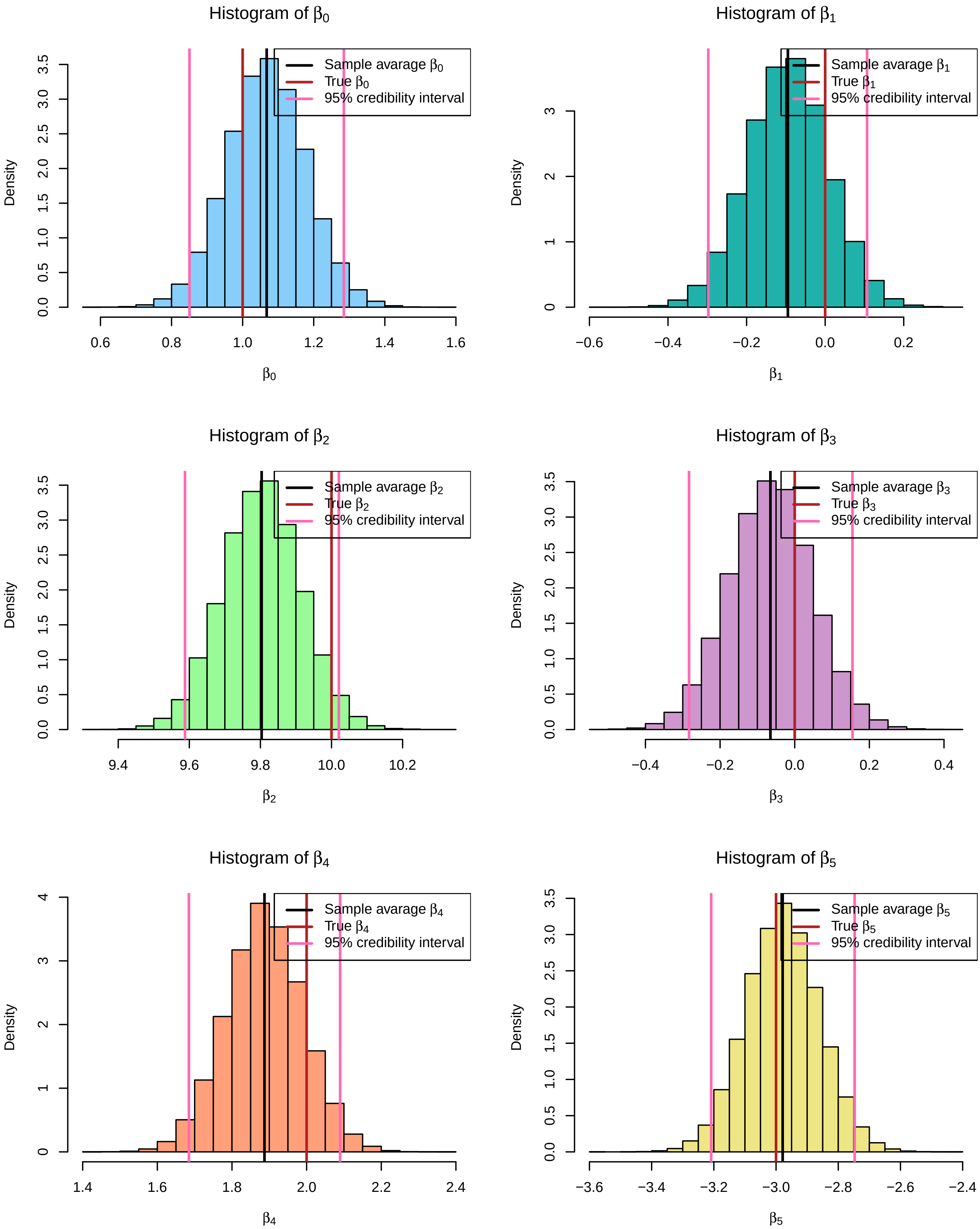
c)

i)



All of the trace plots appear as random scatter, indicating stationarity. This provides evidence for the convergence of the Markov Chains.

ii)



The beta coefficients are approximately normally distributed.

Confidence interval are the relative frequencies of stating valid bounds if you were to re-sample the data. Credibility intervals are the probability that the true parameter is within the stated bounds.

Question 2