

Assignment 1

Abrahams Laylaa (JRGLAY001)

Gumede Sbonelo (GMDSBO006)

Mugeri Khethani (MGRKHE001)

2025-04-14

Question 1

a)

Model formulation for \mathbf{y} :

$$\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \mathbf{e} \text{ where } \mathbf{e} \sim \mathcal{N}(\mathbf{0}, \sigma^2 \mathbf{I}_n)$$

Moments for \mathbf{y} :

$$\mathbb{E}[\mathbf{y}] = \mathbb{E}[\mathbf{X}\boldsymbol{\beta} + \mathbf{e}] = \mathbb{E}[\mathbf{X}\boldsymbol{\beta}] + \mathbb{E}[\mathbf{e}] = \mathbf{X}\boldsymbol{\beta} + \mathbf{0} = \mathbf{X}\boldsymbol{\beta}$$

$$\text{Var}[\mathbf{y}] = \text{Var}[\mathbf{X}\boldsymbol{\beta} + \mathbf{e}] = \text{Var}[\mathbf{X}\boldsymbol{\beta}] + \text{Var}[\mathbf{e}] + 2\text{cov}(\mathbf{X}\boldsymbol{\beta}, \mathbf{e}) = \mathbf{0} + \sigma^2 \mathbf{I}_n + \mathbf{0} = \sigma^2 \mathbf{I}_n$$

Distribution for \mathbf{y} :

$$\mathbf{y} \sim \mathcal{N}(\mathbf{X}\boldsymbol{\beta}, \sigma^2 \mathbf{I}_n)$$

Density for \mathbf{y} :

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

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

Likelihood for $\boldsymbol{\beta}$:

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

Since $\mathbf{M} = \mathbf{M}^{-1} = \mathbf{I}_{k+1}$

Prior distribution for $\boldsymbol{\beta}|\sigma^2$:

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

Prior density for $\boldsymbol{\beta}|\sigma^2$:

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

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

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

Prior distribution for σ^2 :

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

Prior density for σ^2 :

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

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

Assuming that σ^2 is known:

Prior density for $\beta|\sigma^2$:

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

Likelihood for $\beta|\sigma^2$:

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

Posterior for $\beta|\sigma^2$:

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

but $(\mathbf{X}^T \mathbf{X})\hat{\beta} = \mathbf{X}^T \mathbf{y}$

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

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

Completing the square: $\beta^T \mathbf{A}\beta - 2\beta^T \mathbf{b} = (\beta - \mu_\beta)^T \mathbf{A}(\beta - \mu_\beta) - \mu_\beta^T \mathbf{A}\mu_\beta$, where $\mu_\beta = \mathbf{A}^{-1}\mathbf{b}$

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

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

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

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

Therefore:

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

Consider $L(\beta, \mathbf{y}, \mathbf{X})[\beta|\sigma^2]$:

$$L(\beta, \mathbf{y}, \mathbf{X})[\beta|\sigma^2] \propto (\sigma^2)^{-\frac{n}{2}} \exp(-\frac{1}{2\sigma^2}(\mathbf{y}^T \mathbf{y} - 2\beta^T \mathbf{X}^T \mathbf{y} + \beta^T \mathbf{X}^T \mathbf{X} \beta)) (\sigma^2)^{-\frac{k+1}{2}} \exp(-\frac{1}{2\sigma^2}(\beta^T \mathbf{M} \beta - 2\beta^T \mathbf{M} \tilde{\beta} + \tilde{\beta}^T \mathbf{M} \tilde{\beta}))$$

$$L(\beta, \mathbf{y}, \mathbf{X})[\beta|\sigma^2] \propto (\sigma^2)^{-\frac{n+k+1}{2}} \exp(-\frac{1}{2\sigma^2}(\mathbf{y}^T \mathbf{y} + \tilde{\beta}^T \mathbf{M} \tilde{\beta})) \exp(-\frac{1}{2\sigma^2}(\beta^T (\mathbf{M} + \mathbf{X}^T \mathbf{X}) \beta - 2\beta^T ((\mathbf{X}^T \mathbf{X}) \hat{\beta} + \mathbf{M} \tilde{\beta})))$$

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

Completing the square: $\beta^T \mathbf{A} \beta - 2\beta^T \mathbf{b} = (\beta - \mu_\beta)^T \mathbf{A} (\beta - \mu_\beta) - \mu_\beta^T \mathbf{A} \mu_\beta$, where $\mu_\beta = \mathbf{A}^{-1} \mathbf{b}$

$$L(\beta, \mathbf{y}, \mathbf{X})[\beta|\sigma^2] \propto (\sigma^2)^{-\frac{n+k+1}{2}} \exp(-\frac{1}{2\sigma^2}(\mathbf{y}^T \mathbf{y} + \tilde{\beta}^T \mathbf{M} \tilde{\beta})) \exp(-\frac{1}{2\sigma^2}((\beta - \mu_\beta)^T \mathbf{A} (\beta - \mu_\beta) - \mu_\beta^T \mathbf{A} \mu_\beta))$$

$$L(\beta, \mathbf{y}, \mathbf{X})[\beta|\sigma^2] \propto (\sigma^2)^{-\frac{n+k+1}{2}} \exp(-\frac{1}{2\sigma^2}(\mathbf{y}^T \mathbf{y} + \tilde{\beta}^T \mathbf{M} \tilde{\beta} - \mu_\beta^T \mathbf{A} \mu_\beta)) \exp(-\frac{1}{2\sigma^2}(\beta - \mu_\beta)^T \mathbf{A} (\beta - \mu_\beta))$$

Prior density for σ^2 :

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

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

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

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

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

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

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

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

We know that $\det(a\mathbf{A}) = a^k \det(\mathbf{A})$, where \mathbf{A} is a k by k matrix.

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

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

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

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

Therefore:

$$[\sigma^2|\mathbf{y}, \mathbf{X}] \sim \mathcal{IG}(a + \frac{n}{2}, b + \frac{A_2}{2}), \text{ where } a = 1, b = 1 \text{ and } A_2 = \mathbf{y}^T \mathbf{y} + \tilde{\beta}^T \mathbf{M} \tilde{\beta} - \mu_\beta^T (\mathbf{M} + \mathbf{X}^T \mathbf{X}) \mu_\beta$$

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

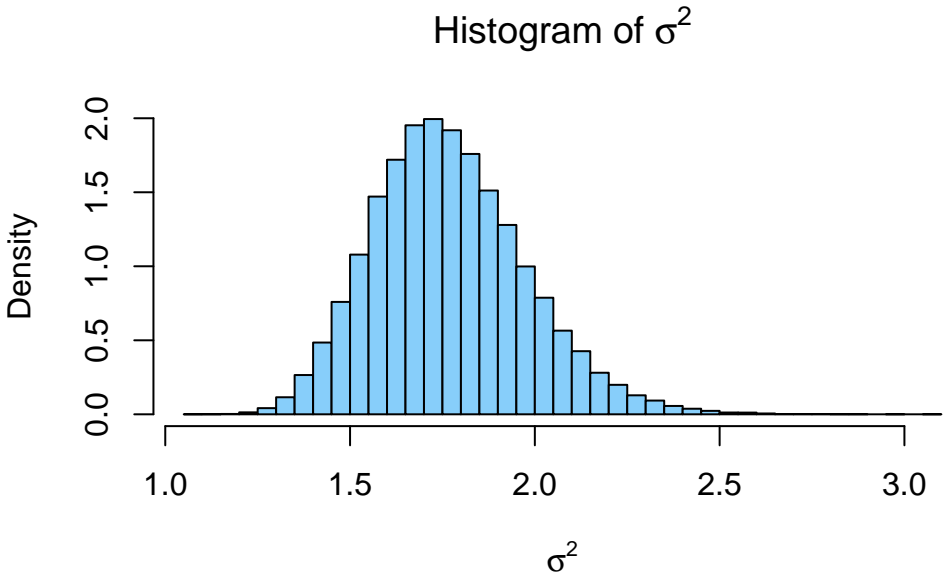
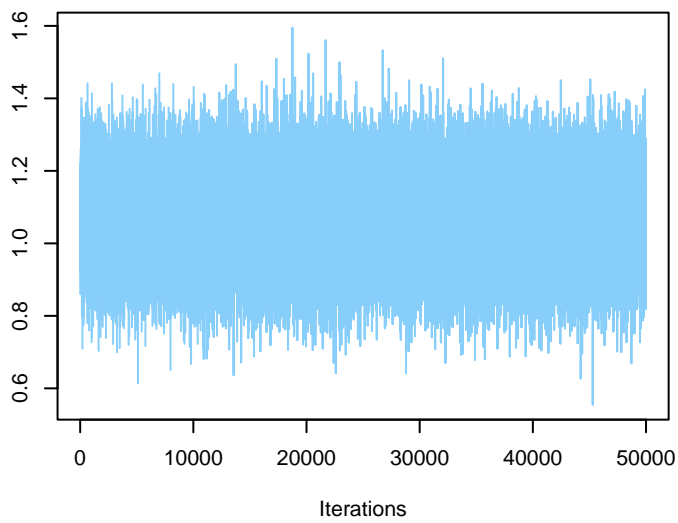
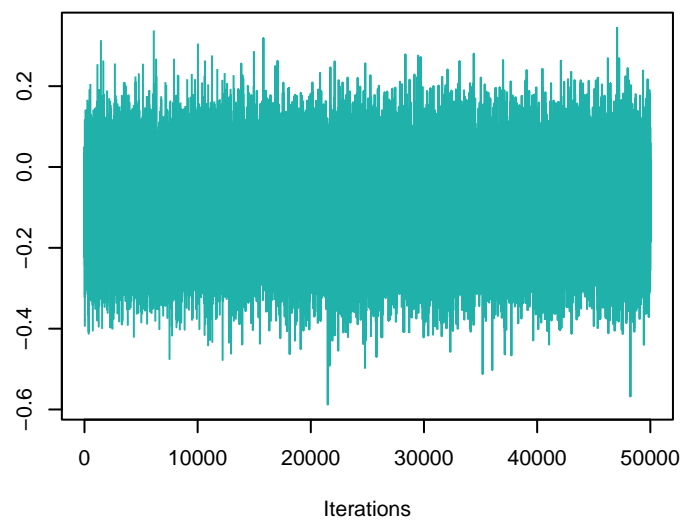
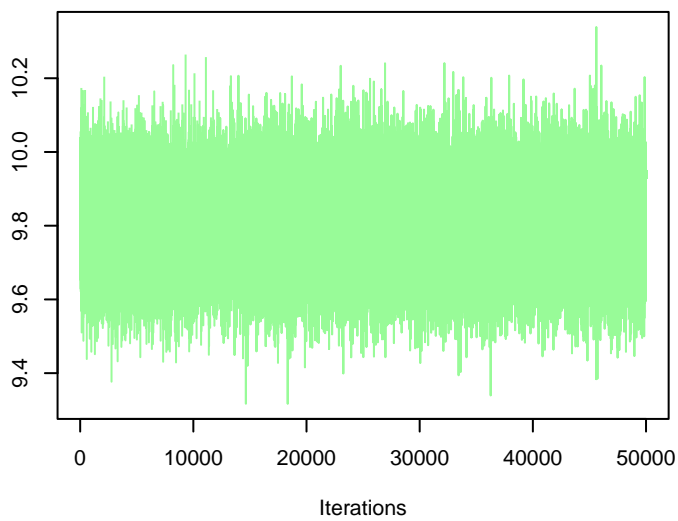
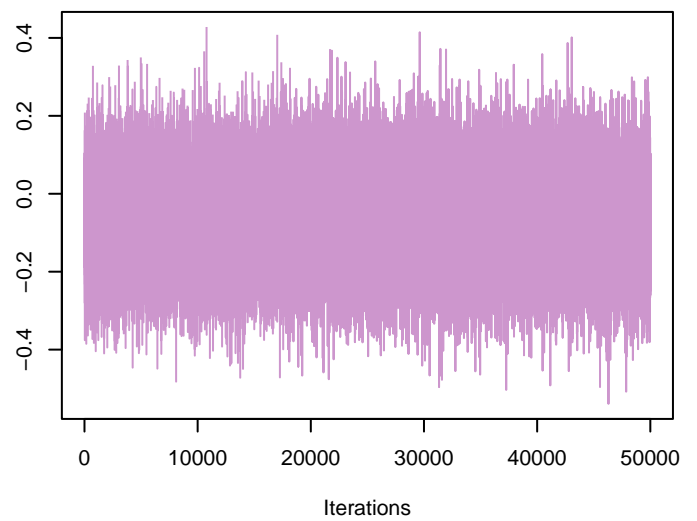
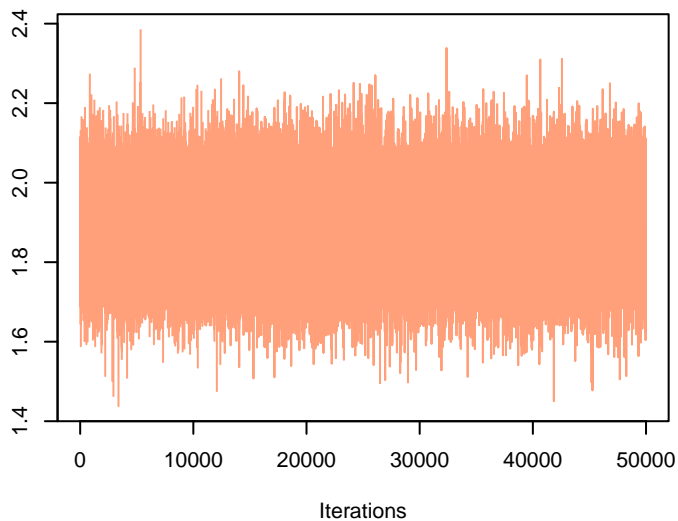
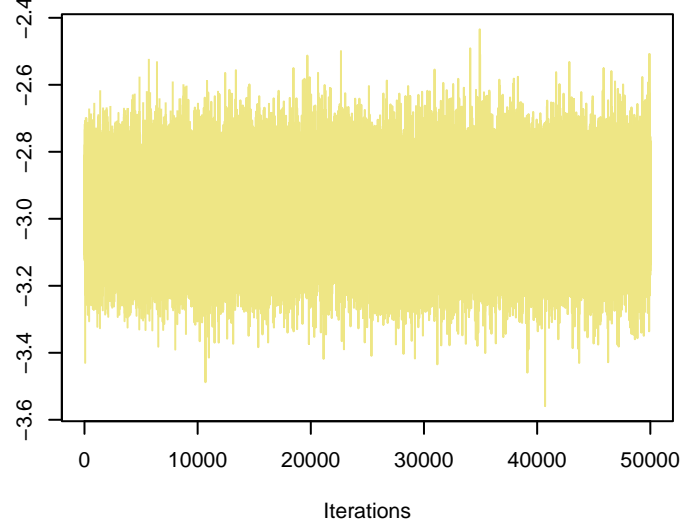


Table 2: First six rows of the β sample values

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

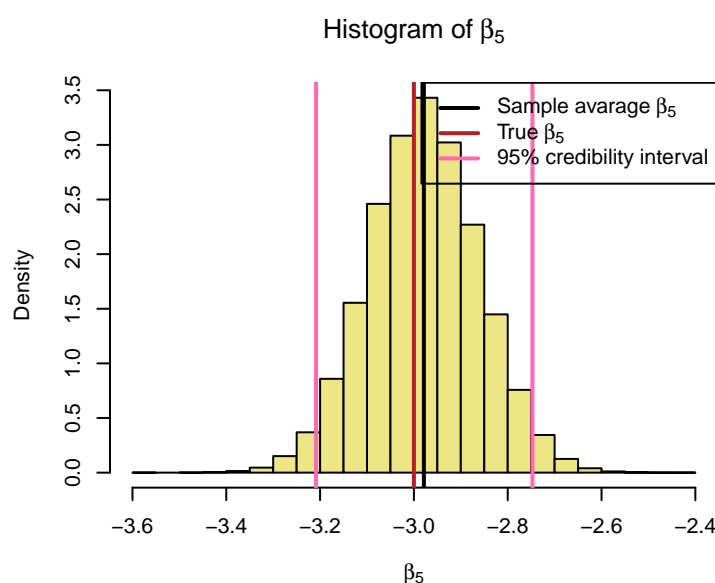
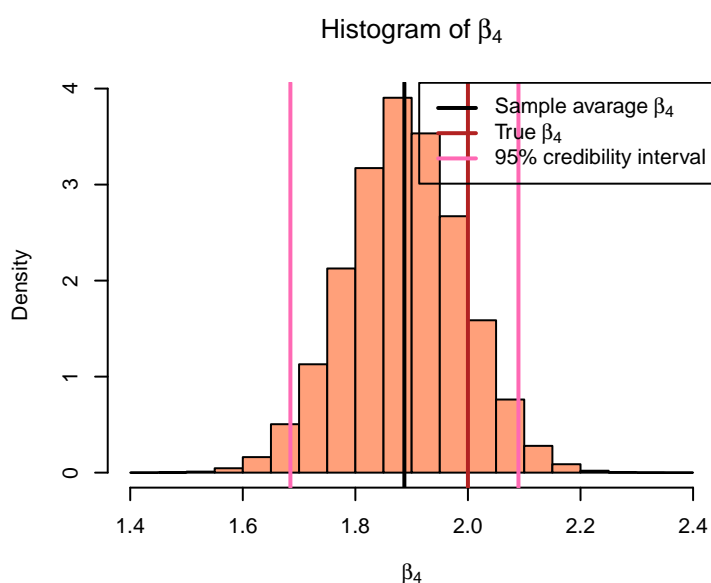
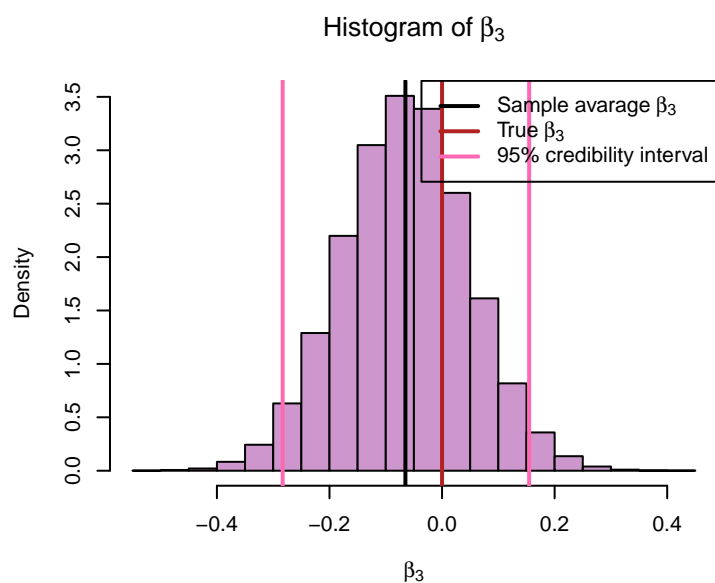
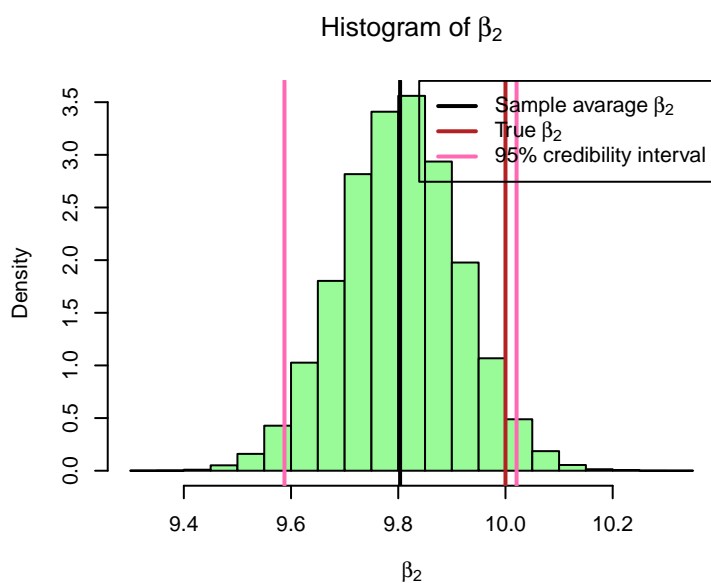
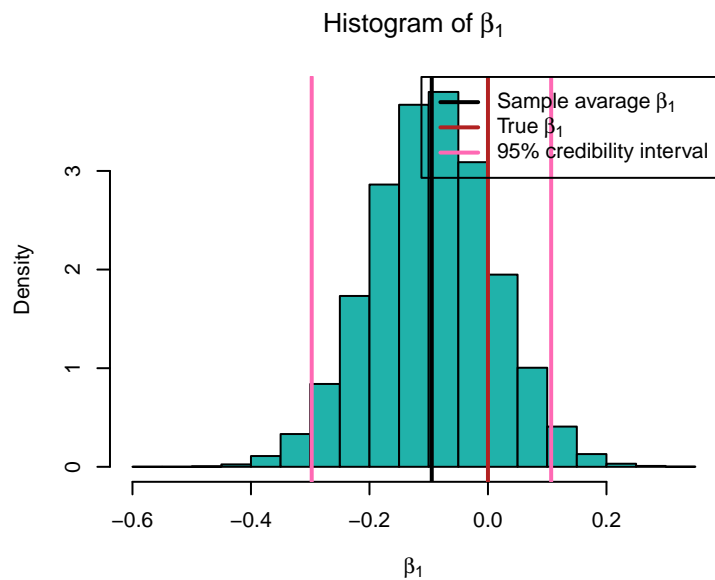
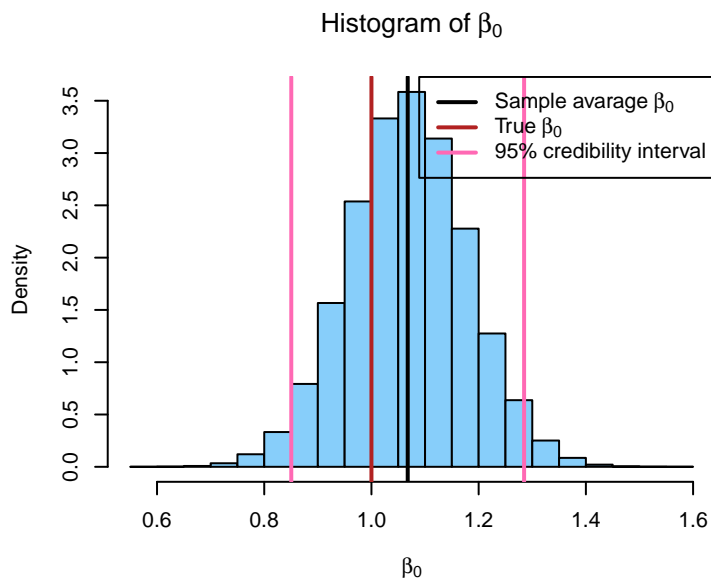
c)

i)

Trace plot of β_0 Trace plot of β_1 Trace plot of β_2 Trace plot of β_3 Trace plot of β_4 Trace plot of β_5 

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. The true beta parameter values should be in the closed interval of these credibility intervals indicated on the plot.

Question 2

a)

Given:

$$Z_i|Y_i \sim Ber(Y_i p_i) = (Y_i p_i)^z (1 - y_i p_i)^{1-z}$$

$$Y_i \sim Ber(\theta_i) = (\theta_i)^y (1 - \theta_i)^{1-y}$$

Derivation of equation 2:

$$\begin{aligned} P(Z_i = 0) &= \sum_{y=0}^1 P(Z_i = 0|Y_i = y)P(Y_i = y) \\ &= (p_i)^0(1 - p_i)^1(\theta_i)^1(1 - \theta_i)^0 + (0p_i)^0(1 - 0p_i)^1(\theta_i)^0(1 - \theta_i)^1 \\ &= (1 - p_i)(\theta_i) + (1 - \theta_i) = 1 - p_i\theta_i \\ P(Y_i = 1) &= (\theta_i)^1(1 - \theta_i)^0 \\ &= \theta_i \end{aligned}$$

$$P(Z_i = 0|Y_i = 1) = (p_i)^0(1 - p_i)^1$$

Therefore :

$$\begin{aligned} P(Y_i = 1|Z_i = 0) &= \frac{P(Z_i=0|Y_i=1)P(Y_i)}{P(Z_i=0)} \\ &= \frac{(1-p_i)\theta_i}{1-p_i\theta_i} \end{aligned}$$

Derivation of equation 4 :

If the fisherman is in cell j then searching cell i gives no information about j , so we are guaranteed to not find the fisherman.

Therefore:

$$\begin{aligned} P(Z_i = 0|Y_j = 1) &= 1 \\ P(Y_j = 1) &= (\theta_j)^1(1 - \theta_j)^0 \\ &= \theta_j \end{aligned}$$

As proved above :

$$P(Z_i = 0) = 1 - p_i\theta_i$$

Therefore:

$$\begin{aligned} \theta_{j,new} &= P(Y_j = 1|Z_i = 0) = \frac{P(Z_i=0|Y_j=1)P(Y_j)}{P(Z_i=0)} \\ \theta_{j,new} &= \frac{\theta_{j,old}}{1-p_i\theta_{i,old}} \end{aligned}$$

b)

Equation 2 is the posterior probability the fisherman is in the cell given that we fail to detect him. Although we don't detect the fisherman he may still be in the cell. So we reduce the probability of occurring in the cell rather than ruling it out. To show this decrease in probability over time we update the new prior (probability of occurrence $\theta_{i,new}$) using the old prior (probability of occurrence $\theta_{i,old}$). Therefore Equation 3 shows how the occurrence probability is adjusted over time as we gain more evidence through the bayesian search.

c)

Approach:

1. Initialize Prior and True Location using Jakes provided functions:

- Generate the initial prior probability distribution using `generate_lost()`.
- Generate the true location of the fisherman using `generate_fisherman()`.
- Store the fisherman's coordinates as `rowf` and `colf`.

2. Variable initialization:

- Create a posterior tracker vector of size 48 to record the posterior probability of the fisherman's true location at each time step.
- Set posterior equal to the initial prior.
- Initialise a boolean called `fishermanfound` to false initially and use this to track whether the fisherman is found or not.

3. Create a for loop which loops through the amount of hours(48):

- Merge the "prior" and detection probability to create a search grid
- Select the cell with the highest probability of successful detection
- If the chosen cell is the fisherman's true location, simulate detection using the detection probability associated with the cell of interest(use `rbinom`).
- Otherwise, detection is automatically set to 0.
- Record the current posterior probability of the true location in `post_tracker[i]`.

4. Update Posterior if Fisherman Not Found:

If the fisherman is not detected use Bayes theorem to update the probability theorem:

- Update the probability in the searched cell using:

$$\theta_{i,\text{new}} = \frac{(1 - p_i)\theta_{i,\text{old}}}{1 - p_i\theta_{i,\text{old}}}$$

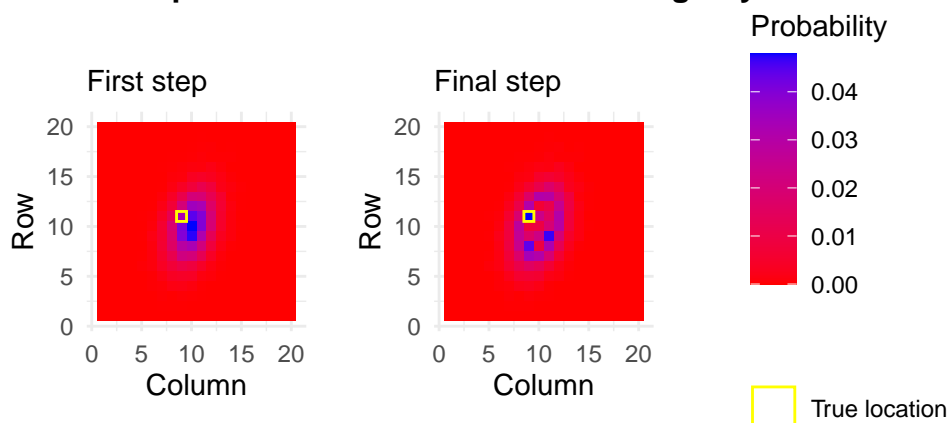
- Update all other cells:

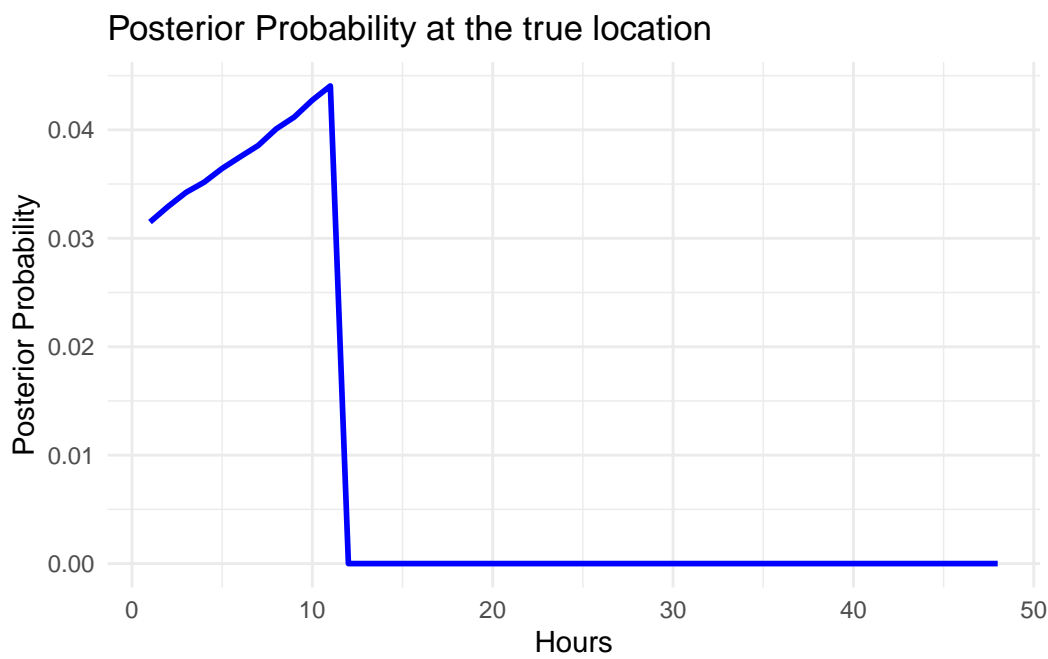
$$\theta_{j,\text{new}} = \frac{\theta_{j,\text{old}}}{1 - p_i\theta_{i,\text{old}}} \quad \text{for } j \neq i$$

5. Successful Detection:

- If detection is successful, break the loop and set the boolean `fisherman found` to true.

Posterior probabilities of occurrence during Bayesian search





d)

If p_i is constant across cells, detection probability no longer varies by location. In this case, the search strategy would focus on cells with the highest prior probability, rather than the highest probability of successful detection. If detection fails we still update the posterior probability using Bayes Theorem and the occurrence probabilities would also still need to be updated. Therefore the only notable change would be which cell we search.

Question 3

Plagiarism Declaration

Abrahams Laylaa (JRGLAY001) declaration

1. I know that plagiarism is a serious form of academic misconduct.
2. I have read the document about avoiding plagiarism, am familiar with its contents and have avoided all forms of plagiarism mentioned there.
3. Where I have used the words of others, I have indicated this by the use of quotation marks.
4. I have referenced all quotations and properly acknowledged ideas borrowed from others.
5. I have not and shall not allow others to plagiarise my work.
6. I declare that this is my own work.
7. I am attaching the summary of the Turnitin match overview (when required to do so).



Signature: Abrahams Laylaa (JRGLAY001)

Gumede Sbonelo (GMDSBO006) declaration

1. I know that plagiarism is a serious form of academic misconduct.
2. I have read the document about avoiding plagiarism, am familiar with its contents and have avoided all forms of plagiarism mentioned there.
3. Where I have used the words of others, I have indicated this by the use of quotation marks.
4. I have referenced all quotations and properly acknowledged ideas borrowed from others.
5. I have not and shall not allow others to plagiarise my work.
6. I declare that this is my own work.
7. I am attaching the summary of the Turnitin match overview (when required to do so).



Signature: Gumede Sbonelo (GMDSBO006)

Mugeri Khethani (MGRKHE001) declaration

1. I know that plagiarism is a serious form of academic misconduct.
2. I have read the document about avoiding plagiarism, am familiar with its contents and have avoided all forms of plagiarism mentioned there.
3. Where I have used the words of others, I have indicated this by the use of quotation marks.
4. I have referenced all quotations and properly acknowledged ideas borrowed from others.
5. I have not and shall not allow others to plagiarise my work.
6. I declare that this is my own work.
7. I am attaching the summary of the Turnitin match overview (when required to do so).



Signature: Mugeri Khethani (MGRKHE001)