Assignment 1

Jorge Abrahams Laylaa (JRGLAY001) Gumede Sbonelo (GMDSBO006)
Mugeri Khethani (MGRKHE001)

2025-04-14

Question 1

(a)

Model formulation for $\mathbf{y}$:

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

Moments for $\mathbf{y}$:

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

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

Distribution for $\mathbf{y}$:

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

Density for $\mathbf{y}$:

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

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

Likelihood for $\boldsymbol{\beta}$:

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

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 $\boldsymbol{\beta}|\sigma^2$:

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

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

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

Prior distribution for σ^2:

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

Prior density for σ^2:

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

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

i.

Assuming that σ^2 is known:
Prior density for $\boldsymbol{\beta}|\sigma^2$:

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

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

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

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

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

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

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

Let $\mathbf{A} = \mathbf{M} + \mathbf{X}^T\mathbf{X}$ and $\mathbf{b} = (\mathbf{X}^T\mathbf{X})\hat{\boldsymbol{\beta}} + \mathbf{M}\tilde{\boldsymbol{\beta}}$
Completing the square: $\boldsymbol{\beta}^T\mathbf{A}\boldsymbol{\beta} - 2\boldsymbol{\beta}^T\mathbf{b} = (\boldsymbol{\beta} - \mu_{\boldsymbol{\beta}})^T\mathbf{A}(\boldsymbol{\beta} - \mu_{\boldsymbol{\beta}}) - \mu_{\boldsymbol{\beta}}^T\mathbf{A}\mu_{\boldsymbol{\beta}}$, where $\mu_{\boldsymbol{\beta}} = \mathbf{A}^{-1}\mathbf{b}$

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

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

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

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

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

ii.

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

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

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

Let $\mathbf{A} = \mathbf{M} + \mathbf{X}^T\mathbf{X}$ and $\mathbf{b} = (\mathbf{X}^T\mathbf{X})\hat{\boldsymbol{\beta}} + \mathbf{M}\tilde{\boldsymbol{\beta}}$
Completing the square: $\boldsymbol{\beta}^T\mathbf{A}\boldsymbol{\beta} - 2\boldsymbol{\beta}^T\mathbf{b} = (\boldsymbol{\beta} - \mu_{\boldsymbol{\beta}})^T\mathbf{A}(\boldsymbol{\beta} - \mu_{\boldsymbol{\beta}}) - \mu_{\boldsymbol{\beta}}^T\mathbf{A}\mu_{\boldsymbol{\beta}}$, where $\mu_{\boldsymbol{\beta}} = \mathbf{A}^{-1}\mathbf{b}$

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

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

Prior density for σ^2:

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

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

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

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

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

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

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

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

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:

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

(b)

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

σ^2
1.968
1.695
1.674
1.982
1.534
1.750

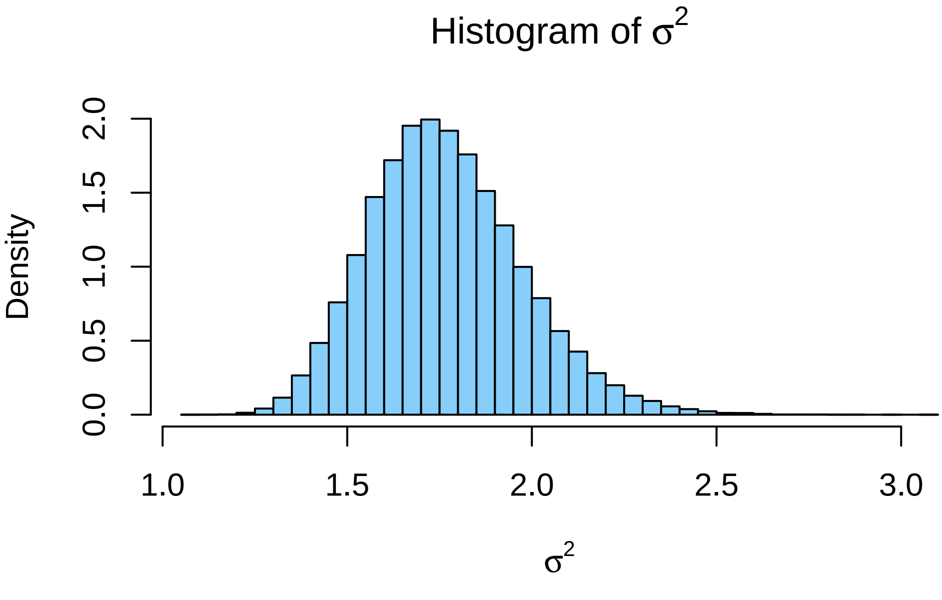
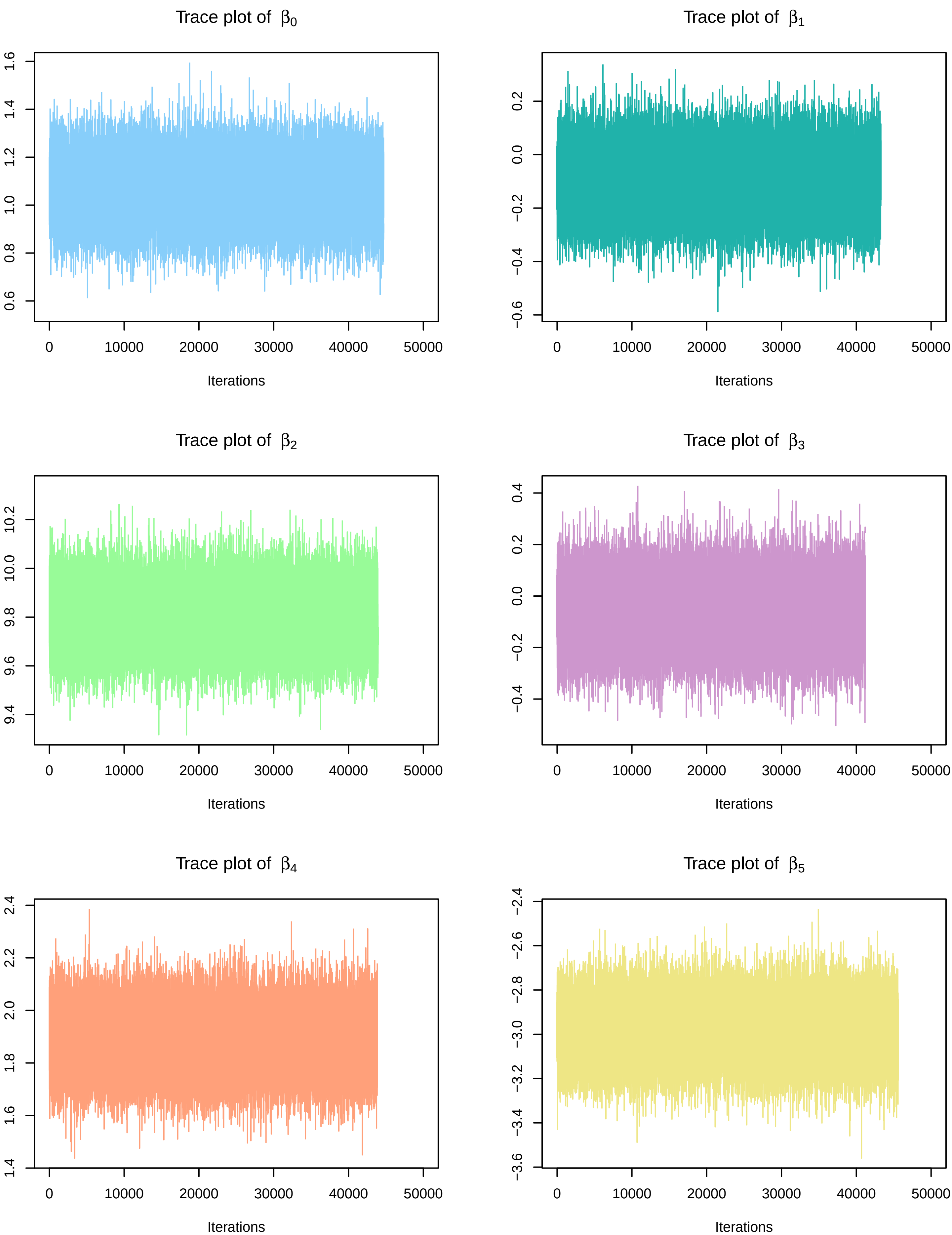


Table 2: First six rows of the β sample values

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

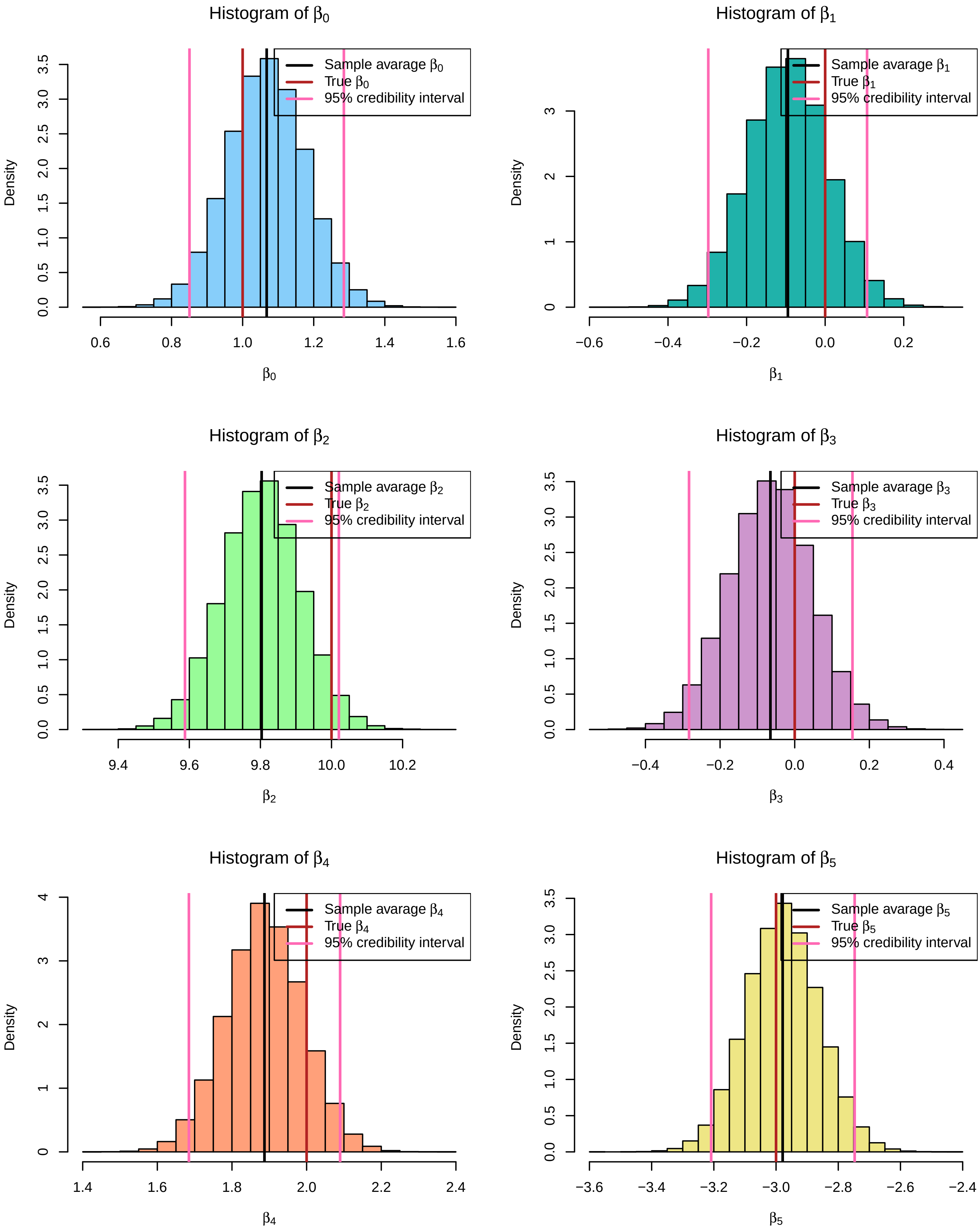
(c)

(i)



All of the trace plots appear as random scatter, indicating stationarity. This provides evidence for the convergence of the Markov Chains. Thus, the sampling quality appears to be good.

(ii)



i.

The posterior distributions of the beta coefficients appear to be approximately normal. However, there seems to be a substantial deviation between the sample mean of the posterior and the true beta coefficient for β_1, β_2, and β_4. This suggests potential bias or high uncertainty in the estimation of these coefficients.

ii.

Confidence intervals represent the relative frequency with which the interval would contain the true parameter if the data were repeatedly resampled. Credible intervals, on the other hand, represent the probability that the true parameter lies within the stated bounds, given the observed data. The true beta parameter values are expected to lie within the credible intervals shown on the plot.

iii.

The credible intervals for β_1 and β_3 include zero, indicating little evidence that these coefficients differ significantly from zero. Therefore, there is weak evidence for retaining these variables in the model. In contrast, the credible intervals for β_0, β_2, β_4, and β_5 do not include zero, providing strong evidence that these variables have a meaningful effect and should be retained in the model.

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:

$P(Z_i = 0) = \sum_{y=0}^{1} P(Z_i = 0|Y_i = y) P(Y_i = y)$

$= (p_i)^0 (1 - pi)^1 (\theta_i)^1 (1 - \theta_i)^0 + (0p_i)^0 (1 - 0p_i)^1 (\theta_i)^0 (1 - \theta_i)^1$

$= (1 - p_i)(\theta_i) + (1 - \theta_i) = 1 - p_i \theta_i$

$P(Y_i = 1) = (\theta_i)^1 (1 - \theta_i)^0$

$= \theta_i$

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

Therefore :

$P(Y_i = 1|Z_i = 0) = \frac{P(Z_i=0|Y_i=1)P(Y_i)}{P(Z_i=0)}$

$= \frac{(1-p_i)\theta_i}{1-p_i\theta_i}$

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:

$P(Z_i = 0|Y_j = 1) = 1$

$P(Y_j = 1) = (\theta_j)^1 (1 - \theta_j)^0$

$= \theta_j$

As proved above :

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

Therefore:

$\theta_{j,new} = P(Y_j = 1|Z_i = 0) = \frac{P(Z_i=0|Y_j=1)P(Y_j)}{P(Z_i=0)}$

$\theta j, new = \frac{\theta_{j,old}}{1-p_i\theta_{i,old}}$

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 occurence$\theta_{i,old}$).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:**

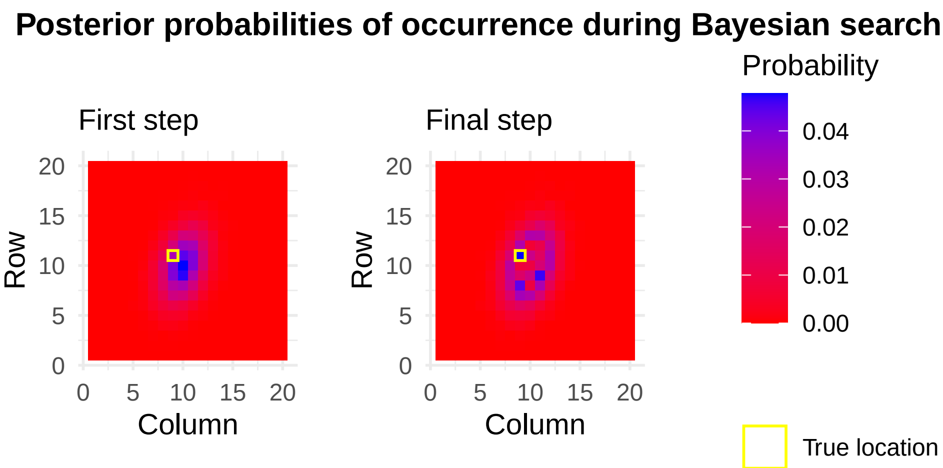
   If the fisherman is not detected use Bayes theorem to update the probability theorem:
   - Update the probability in the searched cell using:

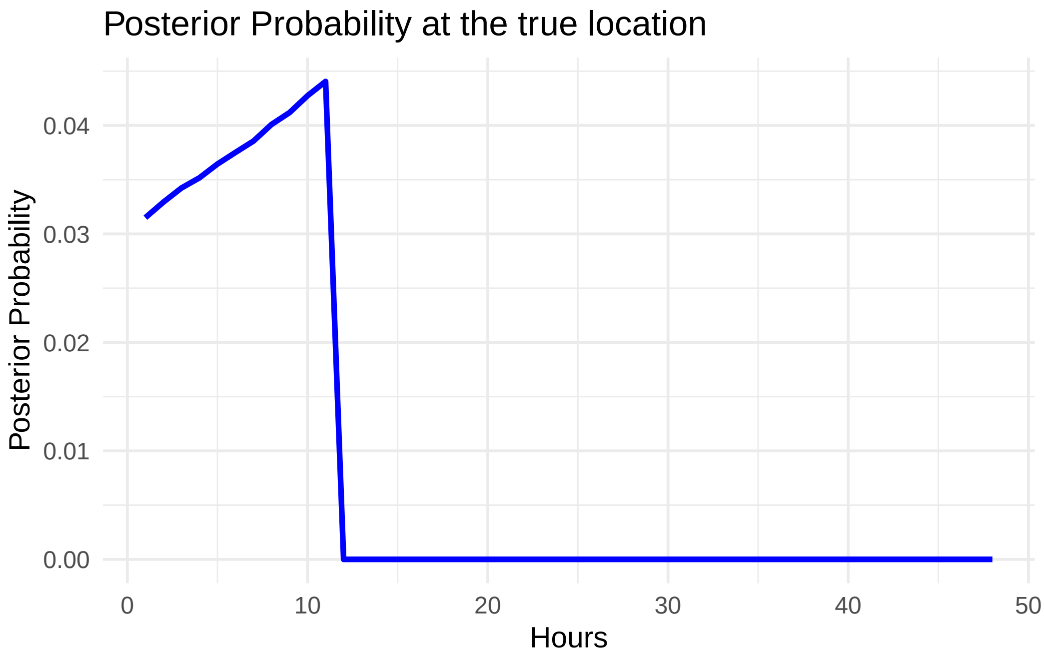
   $$\theta_{i,\text{new}} = \frac{(1-p_i)\theta_{i,\text{old}}}{1-p_i\theta_{i,\text{old}}}$$

   - 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.





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

A Twist on Linear Regressions

Consider the linear regression model where the data is generated as:

$$Y_i = \begin{cases} x_i^\top \beta + e_i, & e_i \sim \mathcal{N}(0, \sigma_1^2), & i \in I_1 \\ x_i^\top \beta + e_i, & e_i \sim \mathcal{N}(0, \sigma_2^2), & i \in I_2 \end{cases}$$

where: (I) is an index set. $(\sigma_1^2 < \sigma_2^2)$

This implies that observations of the second index set have a higher variance than those of the first, potentially representing outliers in the dataset.

As usual, let $\tau_i = \frac{1}{\sigma_i^2}$. If we assume that (I_1) is known, then without loss of generality (w.l.o.g), we can let:

$$I_1 = \{1, \dots, n_1\}, \quad I_2 = \{n_1 + 1, \dots, n\}$$

The likelihood for the data is as follows:

$$L(\mathbf{Y} \mid \beta, \tau_1, \tau_2) = \prod_{i \in I_1} \left(\frac{\tau_1^{1/2}}{\sqrt{2\pi}} \exp\left(-\frac{\tau_1}{2} (Y_i - \mathbf{x}_i^\top \beta)^2 \right) \right) \prod_{i \in I_2} \left(\frac{\tau_2^{1/2}}{\sqrt{2\pi}} \exp\left(-\frac{\tau_2}{2} (Y_i - \mathbf{x}_i^\top \beta)^2 \right) \right)$$

This simplifies to:

$$L(\mathbf{Y} \mid \beta, \tau_1, \tau_2) \propto \tau_1^{n_1/2} \cdot \tau_2^{n_2/2} \cdot \exp\left(-\frac{\tau_1}{2} (\mathbf{Y}_1 - \mathbf{X}_1\beta)^\top (\mathbf{Y}_1 - \mathbf{X}_1\beta) - \frac{\tau_2}{2} (\mathbf{Y}_2 - \mathbf{X}_2\beta)^\top (\mathbf{Y}_2 - \mathbf{X}_2\beta) \right)$$

where:

$$\Sigma^{-1} = \frac{1}{\det(\text{Data Matrix})} \cdot \operatorname{adj}(\text{Data Matrix})$$

$$\Sigma^{-1} = \begin{bmatrix} \tau_1 & 0 & 0 & 0 \\ 0 & \cdots & 0 & 0 \\ 0 & \tau_1 & 0 & 0 \\ 0 & 0 & \tau_2 & 0 \\ 0 & 0 & \cdots & 0 \\ 0 & 0 & 0 & \tau_2 \end{bmatrix}$$

$$\Sigma^{-1} = \tau_1 \underbrace{\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & \cdots & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & \cdots & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}}_{\text{Matrix A}} + \tau_2 \underbrace{\begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & \cdots & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & \cdots & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}}_{\text{Matrix B}}$$

Posterior Derivations

Case 1: Posterior of $\beta \mid \tau_1, \tau_2, X, Y$

Given the following prior distribution:

$\beta \sim \mathcal{N}(\mathbf{0}, \mathbf{T}_0)$

which has the unnormalized form:

$\pi(\beta) \propto \exp\left(-\frac{1}{2}\beta^\top \mathbf{T}_0^{-1}\beta\right)$

The proportional likelihood is given by:

$L(\beta, \mathbf{Y}, \mathbf{X}, \Sigma) \propto \exp\left(-\frac{1}{2}\beta^\top \mathbf{X}^\top \Sigma^{-1}\mathbf{X}\beta\right)$

Therefore, the posterior is:

$\pi(\beta, \tau_1, \tau_2, X, Y) \propto \pi(\beta) \cdot L(\beta, Y, X, \Sigma) \propto \exp\left(-\frac{1}{2}\beta^\top T_0^{-1}\beta\right) \cdot \exp\left(-\frac{1}{2}\beta^\top X^\top \Sigma^{-1} X\beta\right)$

$\mathbf{C} = \mathbf{X}^\top \Sigma^{-1}\mathbf{X} + \mathbf{T}_0^{-1}, \quad \mathbf{D} = \mathbf{X}^\top \Sigma^{-1}\mathbf{Y}$

$\beta^\top \mathbf{C}\beta - 2\beta^\top \mathbf{D} = (\beta - \mu_\beta)^\top \mathbf{C}(\beta - \mu_\beta) - \mu_\beta^\top \mathbf{C}\mu_\beta$

where:

$\mu_\beta = \mathbf{C}^{-1}\mathbf{D}$

The posterior becomes:

$\pi(\beta \mid \tau_1, \tau_2, \mathbf{X}, \mathbf{Y}) \propto \exp\left(-\frac{1}{2}(\beta - \mu_\beta)^\top \mathbf{C}(\beta - \mu_\beta)\right)$

But: $\exp\left(-\frac{1}{2}\mu_B^\top C\mu_B\right) \quad$ is independent of β

Therefore, the posterior simplifies to: $\pi(\beta \mid \tau_1, \tau_2, X, Y) \propto \exp\left(-\frac{1}{2}(\beta - \mu_B)^\top \left(X^\top \Sigma^{-1} X + T_0^{-1}\right)(\beta - \mu_B)\right)$

This is a Gaussian: $\pi(\beta \mid \tau_1, \tau_2, X, Y) \sim \mathcal{N}(\mu_\beta, C^{-1})$

where:

$C^{-1} = \left(X^\top \Sigma^{-1} X + T_0^{-1}\right)^{-1} \quad \mu_\beta = \left(X^\top \Sigma^{-1} Y\right) \times \left(X^\top \Sigma^{-1} X + T_0^{-1}\right)$

For Case 2:) Posterior of $\tau_1|\beta, \tau_2, X, Y$

Given the following prior distribution:$\tau_1 \sim \text{Gamma}(a, b)$

which has the unnormalized form:

$\pi(\tau_1) \propto \tau_1^{a-1} \exp(-b\tau_1)$

The likelihood contribution for group 1 (n1) is:

$L(Y_1, X_1, \beta, \tau_1) \propto \tau_1^{n_1/2} \exp\left[-\frac{\tau_1}{2}(Y_1 - X_1\beta)^\top(Y_1 - X_1\beta)\right].$

Therefore, the posterior is: $\pi(\tau_1 \mid \beta, \tau_2, X, Y) \propto \pi(\tau_1) \cdot L(Y_1, X_1, \beta, \tau_1)$,

$\pi(\tau_1 \mid \beta, \tau_2, X, Y) \propto \tau_1^{a-1} \exp(-b\tau_1) \cdot \tau_1^{n_1/2} \exp\left(-\frac{\tau_1}{2}(Y_1 - X_1\beta)^\top(Y_1 - X_1\beta)\right)$,

$\pi(\tau_1 \mid \beta, \tau_2, X, Y) \propto \tau_1^{a+n_1/2-1} \exp\left(-\tau_1\left(b + \frac{1}{2}(Y_1 - X_1\beta)^\top(Y_1 - X_1\beta)\right)\right).$

The normalized form is: $\pi(\tau_1 \mid \beta, \tau_2, X, Y) \sim \text{Gamma}\left(a + \frac{n_1}{2}, b + \frac{1}{2}(Y_1 - X_1\beta)^\top(Y_1 - X_1\beta)\right)$

Case 3: Posterior of $\tau_2 \mid \tau_1, \beta, X, Y$

Given the following prior distribution:

$\tau_2 \mid \tau_1 \sim \text{Gamma}(a, b) \cdot \mathbb{I}(\tau_2 < \tau_1)$,

which has the unnormalized form:

$\pi(\tau_2 \mid \tau_1) \propto \tau_2^{a-1} \exp(-b\tau_2) \cdot \mathbb{I}(\tau_2 < \tau_1)$

The likelihood contribution for group 2 (n2 = n - n1) is:

$L(Y_2, X_2, \beta, \tau_2) \propto \tau_2^{n_2/2} \exp\left[-\frac{\tau_2}{2}(Y_2 - X_2\beta)^\top(Y_2 - X_2\beta)\right]$

Therefore, the posterior is:

$\pi(\tau_2 \mid \beta, \tau_1, X, Y) \propto \pi(\tau_2 \mid \tau_1) \cdot L(Y_2, X_2, \beta, \tau_2)$,

$\pi(\tau_2 \mid \beta, \tau_1, X, Y) \propto \tau_2^{a-1} \exp(-b\tau_2) \cdot \tau_2^{n_2/2} \exp\left(-\frac{\tau_2}{2}(Y_2 - X_2\beta)^\top(Y_2 - X_2\beta)\right) \cdot \mathbb{I}(\tau_2 < \tau_1)$,

$\pi(\tau_2 \mid \beta, \tau_1, X, Y) \propto \tau_2^{a+n_2/2-1} \exp\left(-\tau_2\left(b + \frac{1}{2}(Y_2 - X_2\beta)^\top(Y_2 - X_2\beta)\right)\right) \cdot \mathbb{I}(\tau_2 < \tau_1)$

The normalized form is: $\pi(\tau_2 \mid \beta, \tau_1, X, Y) \sim \text{Truncated-Gamma}\left(a + \frac{n_2}{2}, b + \frac{1}{2}(Y_2 - X_2\beta)^\top(Y_2 - X_2\beta)\right) \cdot \mathbb{I}(\tau_2 < \tau_1)$

Table 3: First six rows of the β sample values

β_0	β_1	β_2	β_3	β_4	β_5
1.095	0.027	9.679	-0.128	2.168	-2.791
1.115	-0.025	9.844	-0.186	1.868	-3.042
1.114	-0.098	9.756	-0.067	1.810	-3.023
1.093	-0.111	9.769	-0.150	1.820	-3.127
1.135	0.004	9.799	-0.255	1.784	-2.863
1.152	-0.116	10.036	0.084	1.825	-3.091

Table 4: First six rows of τ_1 sample values

τ_1
0.008
0.674
0.866
0.848
0.768
0.827

Table 5: First six rows of τ_2 sample values

τ_2
0.793
1.134
0.947
0.879
0.773
0.687

Plagiarism Declaration

Jorge 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 ideasborrowed 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).

JAbrahams

Signature: Jorge 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 ideasborrowed 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).

SW Gumede

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 ideasborrowed 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).

Mugerika

Signature: Mugeri Khethani (MGRKHE001)