Lab 8: Data Augmentation

Olivier Binette and Rebecca C. Steorts

Friday October 9, 2020

Agenda

- Reminders
- Problem statement
- Go through the lab's tasks
- Office hours

Reminders

- Feel free to send emails for questions or anything
- We're there for you during office ours
  - We can also find other times
- Do let us know if you run into issues, if you have concerns or comments about the class or anything.
  - I'm (Olivier) not a grader and I can bring things up anonymously with Beka or with other TAs.
- The goal of the class is for you to learn and to be prepared for what comes next.

Problem statement

Problem statement

Data points $Y_1, Y_2, \ldots, Y_n$ coming from a **mixture model**:

$$Y_i \sim \sum_{j=1}^{n} w_j N(\mu_j, \varepsilon^2).$$

What does this mean?

For every data point:

- let Z_i be a random variable such that $\mathbb{P}(Z_i = j) = w_j$ for $j = 1, 2, 3,$
- let $Y_i \mid Z_i \sim N(\mu_{Z_i}, \varepsilon^2)$.

Problem statement

Data points $Y_1, Y_2, \dots, Y_n$ coming from a **mixture model**:

$$Y_i \sim \sum_{j=1}^{n} w_j N(\mu_j, \varepsilon^2).$$

What does this mean?

For every data point:

- let Z_i be a random variable such that $\mathbb{P}(Z_i = j) = w_j$ for $j = 1, 2, 3,$
- let $Y_i \mid Z_i \sim N(\mu_{Z_i}, \varepsilon^2)$.

Problem statement

Data points $Y_1, Y_2, \ldots, Y_n$ coming from a **mixture model**:

$$Y_i \sim \sum_{j=1}^{n} w_j N(\mu_j, \varepsilon^2).$$

What does this mean?

For every data point:

- let Z_i be a random variable such that $\mathbb{P}(Z_i = j) = w_j$ for $j = 1, 2, 3,$
- let $Y_i \mid Z_i \sim N(\mu_{Z_i}, \varepsilon^2)$.

Problem statement

Data points $Y_1, Y_2, \ldots, Y_n$ coming from a **mixture model**:

$$Y_i \sim \sum_{j=1}^{n} w_j N(\mu_j, \varepsilon^2).$$

What does this mean?

For every data point:

- let Z_i be a random variable such that $\mathbb{P}(Z_i = j) = w_j$ for $j = 1, 2, 3$,
- let $Y_i \mid Z_i \sim N(\mu_{Z_i}, \varepsilon^2)$.

Problem statement

Data points $Y_1, Y_2, \ldots, Y_n$ coming from a **mixture model**:

$$Y_i \sim \sum_{j=1}^{n} w_j N(\mu_j, \varepsilon^2).$$

What does this mean?

For every data point:

- let Z_i be a random variable such that $\mathbb{P}(Z_i = j) = w_j$ for $j = 1, 2, 3,$
- let $Y_i \mid Z_i \sim N(\mu_{Z_i}, \varepsilon^2)$.

Problem statement

Data points $Y_1, Y_2, \ldots, Y_n$ coming from a **mixture model**:

$$Y_i \sim \sum_{j=1}^{n} w_j N(\mu_j, \varepsilon^2).$$

What does this mean?

For every data point:

- let Z_i be a random variable such that $\mathbb{P}(Z_i = j) = w_j$ for $j = 1, 2, 3$,
- let $Y_i \mid Z_i \sim N(\mu_{Z_i}, \varepsilon^2)$.

Problem statement

In other words:

$$p(Y_i \mid w_{1:3}, \mu_{1:3}, \varepsilon^2) = \sum_{j=1}^{3} w_j N(Y_i; \mu_j, \varepsilon^2)$$

and

$$p(Y_i \mid Z_i, w_{1:3}, \mu_{1:3}, \varepsilon^2) = N(Y_i; \mu_{Z_i}, \varepsilon^2)$$

Problem statement

Let's see what this mixture model could look like in an example.

Let $\mu_1 = -5$, $\mu_2 = 0$ and $\mu_3 = 5$, and let $\varepsilon = 1$. Let $w_j = 1/3$ for $j = 1, 2, 3$.

Now let's generate data from the mixture model:

```
n = 100
mu = c(-5, 0, 5)

Z = sample(1:3, size=n, replace=TRUE)
Y = rnorm(n, mean=mu[Z], sd=1)

hist(Y, breaks=20)
```

Problem statement

Let's see what this mixture model could look like in an example.

Let $\mu_1 = -5$, $\mu_2 = 0$ and $\mu_3 = 5$, and let $\varepsilon = 1$. Let $w_j = 1/3$ for $j = 1, 2, 3$.

Now let's generate data from the mixture model:

```
n = 100
mu = c(-5, 0, 5)

Z = sample(1:3, size=n, replace=TRUE)
Y = rnorm(n, mean=mu[Z], sd=1)

hist(Y, breaks=20)
```

Problem statement

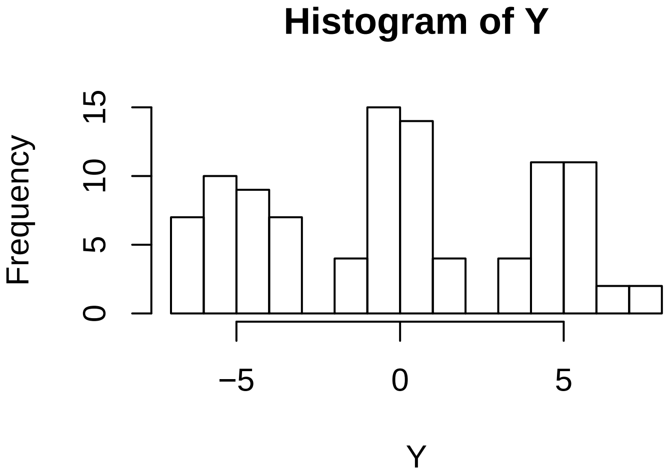
Let's see what this mixture model could look like in an example.

Let $\mu_1 = -5$, $\mu_2 = 0$ and $\mu_3 = 5$, and let $\varepsilon = 1$. Let $w_j = 1/3$ for $j = 1, 2, 3$.

Now let's generate data from the mixture model:

```
n = 100
mu = c(-5, 0, 5)

Z = sample(1:3, size=n, replace=TRUE)
Y = rnorm(n, mean=mu[Z], sd=1)

hist(Y, breaks=20)
```

Problem statement



Problem statement

We have the model

$$Y_i \sim \sum_{j=1}^{n} w_j N(\mu_j, \varepsilon^2),$$

now we need priors on the unknown parameters μ_j, w_j and ε.

Priors

For the means:

$$\mu_j \mid \mu_0, \sigma_0 \sim N(\mu_0, \sigma_0^2)$$
$$\mu_0 \sim N(0, 3)$$
$$\sigma_0^2 \sim IG(2, 2)$$

and recall that $\sigma_0^2 \sim IG(2, 2)$ means that $\phi_0 = 1/\sigma_0^2 \sim \text{Gamma}(2, 2)$.

Problem statement

We have the model

$$Y_i \sim \sum_{j=1}^{n} w_j N(\mu_j, \varepsilon^2),$$

now we need priors on the unknown parameters μ_j, w_j and ε.

Priors

For the means:

$$\mu_j \mid \mu_0, \sigma_0 \sim N(\mu_0, \sigma_0^2)$$
$$\mu_0 \sim N(0, 3)$$
$$\sigma_0^2 \sim IG(2, 2)$$

and recall that $\sigma_0^2 \sim IG(2, 2)$ means that $\phi_0 = 1/\sigma_0^2 \sim \text{Gamma}(2, 2)$.

Problem statement

We have the model

$$Y_i \sim \sum_{j=1}^{n} w_j N(\mu_j, \varepsilon^2),$$

now we need priors on the unknown parameters μ_j, w_j and ε.

Priors

For the means:

$$\mu_j \mid \mu_0, \sigma_0 \sim N(\mu_0, \sigma_0^2)$$
$$\mu_0 \sim N(0, 3)$$
$$\sigma_0^2 \sim IG(2, 2)$$

and recall that $\sigma_0^2 \sim IG(2, 2)$ means that $\phi_0 = 1/\sigma_0^2 \sim \text{Gamma}(2, 2)$.

Problem statement

We have the model

$$Y_i \sim \sum_{j=1}^{n} w_j N(\mu_j, \varepsilon^2),$$

now we need priors on the unknown parameters μ_j, w_j and ε.

Priors

For the means:

$$\begin{aligned} \mu_j \mid \mu_0, \sigma_0 &\sim N(\mu_0, \sigma_0^2) \\ \mu_0 &\sim N(0, 3) \\ \sigma_0^2 &\sim IG(2, 2) \end{aligned}$$

and recall that $\sigma_0^2 \sim IG(2, 2)$ means that $\phi_0 = 1/\sigma_0^2 \sim \text{Gamma}(2, 2)$.

Problem statement

We have the model

$$Y_i \sim \sum_{j=1}^{n} w_j N(\mu_j, \varepsilon^2),$$

now we need priors on the unknown parameters μ_j, w_j and ε.

Priors

For the means:

$$\begin{aligned} \mu_j \mid \mu_0, \sigma_0 &\sim N(\mu_0, \sigma_0^2) \\ \mu_0 &\sim N(0, 3) \\ \sigma_0^2 &\sim IG(2, 2) \end{aligned}$$

and recall that $\sigma_0^2 \sim IG(2,2)$ means that $\phi_0 = 1/\sigma_0^2 \sim \text{Gamma}(2,2)$.

Problem statement

We have the model

$$Y_i \sim \sum_{j=1}^{n} w_j N(\mu_j, \varepsilon^2),$$

now we need priors on the unknown parameters μ_j, w_j and ε.

Priors

For the means:

$$\begin{aligned} \mu_j \mid \mu_0, \sigma_0 &\sim N(\mu_0, \sigma_0^2) \\ \mu_0 &\sim N(0, 3) \\ \sigma_0^2 &\sim IG(2, 2) \end{aligned}$$

and recall that $\sigma_0^2 \sim IG(2, 2)$ means that $\phi_0 = 1/\sigma_0^2 \sim \text{Gamma}(2, 2)$.

Problem statement

Priors

For the mixture weights:

$$(w_1, w_2, w_3) \sim \text{Dirichlet}(\mathbf{1})$$

which means that $p(w_1, w_2, w_3) \propto 1$.

Recall that, in general,

$$\begin{aligned}(w_1, w_2, w_3) &\sim \text{Dirichlet}(\alpha_1, \alpha_2, \alpha_3) \\ \Rightarrow p(w_1, w_2, w_3) &\propto w_1^{\alpha_1 - 1} w_2^{\alpha_2 - 1} w_3^{\alpha_3 - 1}.\end{aligned}$$

Problem statement

Priors

For the mixture weights:

$$(w_1, w_2, w_3) \sim \text{Dirichlet}(\mathbf{1})$$

which means that $p(w_1, w_2, w_3) \propto 1$.

Recall that, in general,

$$(w_1, w_2, w_3) \sim \text{Dirichlet}(\alpha_1, \alpha_2, \alpha_3)$$
$$\Rightarrow p(w_1, w_2, w_3) \propto w_1^{\alpha_1 - 1} w_2^{\alpha_2 - 1} w_3^{\alpha_3 - 1}.$$

Problem statement

Priors

For the mixture weights:

$$(w_1, w_2, w_3) \sim \text{Dirichlet}(\mathbf{1})$$

which means that $p(w_1, w_2, w_3) \propto 1$.

Recall that, in general,

$$\begin{aligned} &(w_1, w_2, w_3) \sim \text{Dirichlet}(\alpha_1, \alpha_2, \alpha_3) \\ \Rightarrow\, &p(w_1, w_2, w_3) \propto w_1^{\alpha_1 - 1} w_2^{\alpha_2 - 1} w_3^{\alpha_3 - 1}. \end{aligned}$$

Problem statement

Priors

For the variance:

$$\varepsilon^2 \sim IG(2, 2)$$

This means that

$$\tau = 1/\varepsilon^2 \sim \text{Gamma}(2, 2)$$

Problem statement

Priors

For the variance:

$$\varepsilon^2 \sim IG(2, 2)$$

This means that

$$\tau = 1/\varepsilon^2 \sim \text{Gamma}(2, 2)$$

Problem statement

In summary,

$$
\begin{aligned}
p(Y_i|\mu_1,\mu_2,\mu_3,w_1,w_2,w_3,\tau) &= \sum_{j=1}^{3} w_i N(\mu_j,\tau^{-1}) \\
\mu_j|\mu_0,\sigma_0^2 &\sim N(\mu_0,\phi_0^{-1}) \\
\mu_0 &\sim N(0,3) \\
\phi_0 &\sim \text{Gamma}(2,2) \\
(w_1,w_2,w_3) &\sim \textit{Dirichlet}(\mathbf{1}) \\
\tau &\sim \text{Gamma}(2,2),
\end{aligned}
$$

for $i = 1, \ldots n$.

Task 1

Derive the joint posterior $p(w_1, w_2, w_3, \mu_1, \mu_2, \mu_3, \varepsilon^2, \mu_0, \sigma_0^2 | Y_{1:N})$ up to a normalizing constant.

Let's do the derivations using $\tau = 1/\varepsilon^2$ and $\phi_0 = 1/\sigma_0^2$.

The posterior distribution is always proportional to the full joint distribution:

$$
\begin{aligned}
&p(Y_{1:n}, \mu_{1:3}, \mu_0, \phi_0, w_{1:3}, \tau) \\
=&p(Y_{1:n} \mid \mu_{1:3}, w_{1:3}, \tau) p(\mu_{1:3}, \mu_0, \phi_0, w_{1:3}, \tau) \\
=&p(Y_{1:n} \mid \mu_{1:3}, w_{1:3}, \tau) p(\mu_{1:3} \mid \mu_0, \phi_0) p(\mu_0) p(\phi_0) p(w_{1:3}) p(\tau) \\
=&p(Y_{1:n} \mid \mu_{1:3}, w_{1:3}, \tau) p(\mu_{1:3} \mid \mu_0, \phi_0) p(\mu_0) p(\phi_0) p(\tau) \\
=&\left(\prod_{i=1}^{n} p(Y_i \mid \mu_{1:3}, w_{1:3}, \tau) \right) \left(\prod_{j=1}^{3} p(\mu_j \mid \mu_0, \phi_0) \right) p(\mu_0) p(\phi_0) p(\tau).
\end{aligned}
$$

Task 1

Derive the joint posterior $p(w_1, w_2, w_3, \mu_1, \mu_2, \mu_3, \varepsilon^2, \mu_0, \sigma_0^2 | Y_{1:N})$ up to a normalizing constant.

Let's do the derivations using $\tau = 1/\varepsilon^2$ and $\phi_0 = 1/\sigma_0^2$.

The posterior distribution is always proportional to the full joint distribution:

$$
\begin{aligned}
& p(Y_{1:n}, \mu_{1:3}, \mu_0, \phi_0, w_{1:3}, \tau) \\
=& p(Y_{1:n} \mid \mu_{1:3}, w_{1:3}, \tau) p(\mu_{1:3}, \mu_0, \phi_0, w_{1:3}, \tau) \\
=& p(Y_{1:n} \mid \mu_{1:3}, w_{1:3}, \tau) p(\mu_{1:3} \mid \mu_0, \phi_0) p(\mu_0) p(\phi_0) p(w_{1:3}) p(\tau) \\
=& p(Y_{1:n} \mid \mu_{1:3}, w_{1:3}, \tau) p(\mu_{1:3} \mid \mu_0, \phi_0) p(\mu_0) p(\phi_0) p(\tau) \\
=& \left(\prod_{i=1}^{n} p(Y_i \mid \mu_{1:3}, w_{1:3}, \tau) \right) \left(\prod_{j=1}^{3} p(\mu_j \mid \mu_0, \phi_0) \right) p(\mu_0) p(\phi_0) p(\tau).
\end{aligned}
$$

Task 1

Derive the joint posterior $p(w_1, w_2, w_3, \mu_1, \mu_2, \mu_3, \varepsilon^2, \mu_0, \sigma_0^2 | Y_{1:N})$ up to a normalizing constant.

Let's do the derivations using $\tau = 1/\varepsilon^2$ and $\phi_0 = 1/\sigma_0^2$.

The posterior distribution is always proportional to the full joint distribution:

$$
\begin{aligned}
& p(Y_{1:n}, \mu_{1:3}, \mu_0, \phi_0, w_{1:3}, \tau) \\
=& p(Y_{1:n} \mid \mu_{1:3}, w_{1:3}, \tau) p(\mu_{1:3}, \mu_0, \phi_0, w_{1:3}, \tau) \\
=& p(Y_{1:n} \mid \mu_{1:3}, w_{1:3}, \tau) p(\mu_{1:3} \mid \mu_0, \phi_0) p(\mu_0) p(\phi_0) p(w_{1:3}) p(\tau) \\
=& p(Y_{1:n} \mid \mu_{1:3}, w_{1:3}, \tau) p(\mu_{1:3} \mid \mu_0, \phi_0) p(\mu_0) p(\phi_0) p(\tau) \\
=& \left(\prod_{i=1}^{n} p(Y_i \mid \mu_{1:3}, w_{1:3}, \tau) \right) \left(\prod_{j=1}^{3} p(\mu_j \mid \mu_0, \phi_0) \right) p(\mu_0) p(\phi_0) p(\tau).
\end{aligned}
$$

Task 1

Derive the joint posterior $p(w_1, w_2, w_3, \mu_1, \mu_2, \mu_3, \varepsilon^2, \mu_0, \sigma_0^2 | Y_{1:N})$ up to a normalizing constant.

Let's do the derivations using $\tau = 1/\varepsilon^2$ and $\phi_0 = 1/\sigma_0^2$.

The posterior distribution is always proportional to the full joint distribution:

$$
\begin{aligned}
&p(Y_{1:n}, \mu_{1:3}, \mu_0, \phi_0, w_{1:3}, \tau) \\
=&p(Y_{1:n} \mid \mu_{1:3}, w_{1:3}, \tau) p(\mu_{1:3}, \mu_0, \phi_0, w_{1:3}, \tau) \\
=&p(Y_{1:n} \mid \mu_{1:3}, w_{1:3}, \tau) p(\mu_{1:3} \mid \mu_0, \phi_0) p(\mu_0) p(\phi_0) p(w_{1:3}) p(\tau) \\
=&p(Y_{1:n} \mid \mu_{1:3}, w_{1:3}, \tau) p(\mu_{1:3} \mid \mu_0, \phi_0) p(\mu_0) p(\phi_0) p(\tau) \\
=&\left(\prod_{i=1}^{n} p(Y_i \mid \mu_{1:3}, w_{1:3}, \tau) \right) \left(\prod_{j=1}^{3} p(\mu_j \mid \mu_0, \phi_0) \right) p(\mu_0) p(\phi_0) p(\tau).
\end{aligned}
$$

Task 1

Derive the joint posterior $p(w_1, w_2, w_3, \mu_1, \mu_2, \mu_3, \varepsilon^2, \mu_0, \sigma_0^2 | Y_{1:N})$ up to a normalizing constant.

Let's do the derivations using $\tau = 1/\varepsilon^2$ and $\phi_0 = 1/\sigma_0^2$.

The posterior distribution is always proportional to the full joint distribution:

$$
\begin{aligned}
& p(Y_{1:n}, \mu_{1:3}, \mu_0, \phi_0, w_{1:3}, \tau) \\
=& p(Y_{1:n} \mid \mu_{1:3}, w_{1:3}, \tau) p(\mu_{1:3}, \mu_0, \phi_0, w_{1:3}, \tau) \\
=& p(Y_{1:n} \mid \mu_{1:3}, w_{1:3}, \tau) p(\mu_{1:3} \mid \mu_0, \phi_0) p(\mu_0) p(\phi_0) p(w_{1:3}) p(\tau) \\
=& p(Y_{1:n} \mid \mu_{1:3}, w_{1:3}, \tau) p(\mu_{1:3} \mid \mu_0, \phi_0) p(\mu_0) p(\phi_0) p(\tau) \\
=& \left(\prod_{i=1}^{n} p(Y_i \mid \mu_{1:3}, w_{1:3}, \tau) \right) \left(\prod_{j=1}^{3} p(\mu_j \mid \mu_0, \phi_0) \right) p(\mu_0) p(\phi_0) p(\tau).
\end{aligned}
$$

Task 1

Derive the joint posterior $p(w_1, w_2, w_3, \mu_1, \mu_2, \mu_3, \varepsilon^2, \mu_0, \sigma_0^2 | Y_{1:N})$ up to a normalizing constant.

Let's do the derivations using $\tau = 1/\varepsilon^2$ and $\phi_0 = 1/\sigma_0^2$.

The posterior distribution is always proportional to the full joint distribution:

$$
\begin{aligned}
&p(Y_{1:n}, \mu_{1:3}, \mu_0, \phi_0, w_{1:3}, \tau)\\
=&p(Y_{1:n} \mid \mu_{1:3}, w_{1:3}, \tau) p(\mu_{1:3}, \mu_0, \phi_0, w_{1:3}, \tau)\\
=&p(Y_{1:n} \mid \mu_{1:3}, w_{1:3}, \tau) p(\mu_{1:3} \mid \mu_0, \phi_0) p(\mu_0) p(\phi_0) p(w_{1:3}) p(\tau)\\
=&p(Y_{1:n} \mid \mu_{1:3}, w_{1:3}, \tau) p(\mu_{1:3} \mid \mu_0, \phi_0) p(\mu_0) p(\phi_0) p(\tau)\\
=&\left(\prod_{i=1}^{n} p(Y_i \mid \mu_{1:3}, w_{1:3}, \tau)\right) \left(\prod_{j=1}^{3} p(\mu_j \mid \mu_0, \phi_0)\right) p(\mu_0) p(\phi_0) p(\tau).
\end{aligned}
$$

Task 1

Derive the joint posterior $p(w_1, w_2, w_3, \mu_1, \mu_2, \mu_3, \varepsilon^2, \mu_0, \sigma_0^2 | Y_{1:N})$ up to a normalizing constant.

Let's do the derivations using $\tau = 1/\varepsilon^2$ and $\phi_0 = 1/\sigma_0^2$.

The posterior distribution is always proportional to the full joint distribution:

$$
\begin{aligned}
&p(Y_{1:n}, \mu_{1:3}, \mu_0, \phi_0, w_{1:3}, \tau) \\
=&p(Y_{1:n} \mid \mu_{1:3}, w_{1:3}, \tau) p(\mu_{1:3}, \mu_0, \phi_0, w_{1:3}, \tau) \\
=&p(Y_{1:n} \mid \mu_{1:3}, w_{1:3}, \tau) p(\mu_{1:3} \mid \mu_0, \phi_0) p(\mu_0) p(\phi_0) p(w_{1:3}) p(\tau) \\
=&p(Y_{1:n} \mid \mu_{1:3}, w_{1:3}, \tau) p(\mu_{1:3} \mid \mu_0, \phi_0) p(\mu_0) p(\phi_0) p(\tau) \\
=&\left(\prod_{i=1}^{n} p(Y_i \mid \mu_{1:3}, w_{1:3}, \tau)\right)\left(\prod_{j=1}^{3} p(\mu_j \mid \mu_0, \phi_0)\right) p(\mu_0) p(\phi_0) p(\tau).
\end{aligned}
$$

Task 1

Derive the joint posterior $p(w_1, w_2, w_3, \mu_1, \mu_2, \mu_3, \varepsilon^2, \mu_0, \sigma_0^2 | Y_{1:N})$ up to a normalizing constant.

Let's do the derivations using $\tau = 1/\varepsilon^2$ and $\phi_0 = 1/\sigma_0^2$.

The posterior distribution is always proportional to the full joint distribution:

$$
\begin{aligned}
&p(Y_{1:n}, \mu_{1:3}, \mu_0, \phi_0, w_{1:3}, \tau) \\
=&p(Y_{1:n} \mid \mu_{1:3}, w_{1:3}, \tau) p(\mu_{1:3}, \mu_0, \phi_0, w_{1:3}, \tau) \\
=&p(Y_{1:n} \mid \mu_{1:3}, w_{1:3}, \tau) p(\mu_{1:3} \mid \mu_0, \phi_0) p(\mu_0) p(\phi_0) p(w_{1:3}) p(\tau) \\
=&p(Y_{1:n} \mid \mu_{1:3}, w_{1:3}, \tau) p(\mu_{1:3} \mid \mu_0, \phi_0) p(\mu_0) p(\phi_0) p(\tau) \\
=&\left(\prod_{i=1}^{n} p(Y_i \mid \mu_{1:3}, w_{1:3}, \tau)\right) \left(\prod_{j=1}^{3} p(\mu_j \mid \mu_0, \phi_0)\right) p(\mu_0) p(\phi_0) p(\tau).
\end{aligned}
$$

Task 1

The full joint:

$$\left(\prod_{i=1}^{n} p(Y_i \mid \mu_{1:3}, w_{1:3}, \tau)\right)\left(\prod_{j=1}^{3} p(\mu_j \mid \mu_0, \phi_0)\right) p(\mu_0)p(\phi_0)p(\tau).$$

And

$$p(Y_i \mid \mu_{1:3}, w_{1:3}, \tau) = \sum_{j=1}^{3} w_j N(Y_i; \mu_j, \tau),$$

$$p(\mu_j \mid \mu_0, \phi_0) = N(\mu_j; \mu_0, \phi_0^{-1}),$$

$$p(\mu_0) = N(\mu_0; 0, 3),$$

$$p(\phi_0) = \mathrm{Gamma}(\phi_0; 2, 2),$$

$$p(\tau) = \mathrm{Gamma}(\tau; 2, 2).$$

Task 1

The full joint:

$$\left(\prod_{i=1}^{n} p(Y_i \mid \mu_{1:3}, w_{1:3}, \tau)\right)\left(\prod_{j=1}^{3} p(\mu_j \mid \mu_0, \phi_0)\right) p(\mu_0)p(\phi_0)p(\tau).$$

And

$$p(Y_i \mid \mu_{1:3}, w_{1:3}, \tau) = \sum_{j=1}^{3} w_j N(Y_i; \mu_j, \tau),$$

$$p(\mu_j \mid \mu_0, \phi_0) = N(\mu_j; \mu_0, \phi_0^{-1}),$$

$$p(\mu_0) = N(\mu_0; 0, 3),$$

$$p(\phi_0) = \text{Gamma}(\phi_0; 2, 2),$$

$$p(\tau) = \text{Gamma}(\tau; 2, 2).$$

Task 1

The full joint:

$$\left(\prod_{i=1}^{n} p(Y_i \mid \mu_{1:3}, w_{1:3}, \tau)\right)\left(\prod_{j=1}^{3} p(\mu_j \mid \mu_0, \phi_0)\right) p(\mu_0)p(\phi_0)p(\tau).$$

And

$$p(Y_i \mid \mu_{1:3}, w_{1:3}, \tau) = \sum_{j=1}^{3} w_j N(Y_i; \mu_j, \tau),$$

$$p(\mu_j \mid \mu_0, \phi_0) = N(\mu_j; \mu_0, \phi_0^{-1}),$$

$$p(\mu_0) = N(\mu_0; 0, 3),$$

$$p(\phi_0) = \text{Gamma}(\phi_0; 2, 2),$$

$$p(\tau) = \text{Gamma}(\tau; 2, 2).$$

Task 1

The full joint:

$$\left(\prod_{i=1}^{n} p(Y_i \mid \mu_{1:3}, w_{1:3}, \tau)\right)\left(\prod_{j=1}^{3} p(\mu_j \mid \mu_0, \phi_0)\right) p(\mu_0)p(\phi_0)p(\tau).$$

And

$$p(Y_i \mid \mu_{1:3}, w_{1:3}, \tau) = \sum_{j=1}^{3} w_j N(Y_i; \mu_j, \tau),$$

$$p(\mu_j \mid \mu_0, \phi_0) = N(\mu_j; \mu_0, \phi_0^{-1}),$$

$$p(\mu_0) = N(\mu_0; 0, 3),$$

$$p(\phi_0) = \text{Gamma}(\phi_0; 2, 2),$$

$$p(\tau) = \text{Gamma}(\tau; 2, 2).$$

Task 1

The full joint:

$$\left(\prod_{i=1}^{n} p(Y_i \mid \mu_{1:3}, w_{1:3}, \tau)\right)\left(\prod_{j=1}^{3} p(\mu_j \mid \mu_0, \phi_0)\right) p(\mu_0)p(\phi_0)p(\tau).$$

And

$$p(Y_i \mid \mu_{1:3}, w_{1:3}, \tau) = \sum_{j=1}^{3} w_j N(Y_i; \mu_j, \tau),$$

$$p(\mu_j \mid \mu_0, \phi_0) = N(\mu_j; \mu_0, \phi_0^{-1}),$$

$$p(\mu_0) = N(\mu_0; 0, 3),$$

$$p(\phi_0) = \text{Gamma}(\phi_0; 2, 2),$$

$$p(\tau) = \text{Gamma}(\tau; 2, 2).$$

Task 1

The full joint:

$$\left(\prod_{i=1}^{n} p(Y_i \mid \mu_{1:3}, w_{1:3}, \tau)\right)\left(\prod_{j=1}^{3} p(\mu_j \mid \mu_0, \phi_0)\right) p(\mu_0)p(\phi_0)p(\tau).$$

And

$$p(Y_i \mid \mu_{1:3}, w_{1:3}, \tau) = \sum_{j=1}^{3} w_j N(Y_i; \mu_j, \tau),$$

$$p(\mu_j \mid \mu_0, \phi_0) = N(\mu_j; \mu_0, \phi_0^{-1}),$$

$$p(\mu_0) = N(\mu_0; 0, 3),$$

$$p(\phi_0) = \text{Gamma}(\phi_0; 2, 2),$$

$$p(\tau) = \text{Gamma}(\tau; 2, 2).$$

Task 2

Derive the full conditionals for all the parameters up to a normalizing constant.

$$
\begin{aligned}
&p(w_{1:3} \mid Y_{1:n}, \mu_{1:3}, \mu_0, \phi_0, \tau) \\
&\propto \left(\prod_{i=1}^{n} p(Y_i \mid \mu_{1:3}, w_{1:3}, \tau) \right) \left(\prod_{j=1}^{3} p(\mu_j \mid \mu_0, \phi_0) \right) p(\mu_0) p(\phi_0) p(\tau) \\
&\propto \left(\prod_{i=1}^{n} p(Y_i \mid \mu_{1:3}, w_{1:3}, \tau) \right) \\
&\propto \prod_{i=1}^{n} \left(\sum_{j=1}^{3} w_j N(Y_i; \mu_j, \tau) \right)
\end{aligned}
$$

Task 2

Derive the full conditionals for all the parameters up to a normalizing constant.

$$
\begin{aligned}
& p(w_{1:3} \mid Y_{1:n}, \mu_{1:3}, \mu_0, \phi_0, \tau) \\
\propto & \left(\prod_{i=1}^{n} p(Y_i \mid \mu_{1:3}, w_{1:3}, \tau)\right) \left(\prod_{j=1}^{3} p(\mu_j \mid \mu_0, \phi_0)\right) p(\mu_0) p(\phi_0) p(\tau) \\
\propto & \left(\prod_{i=1}^{n} p(Y_i \mid \mu_{1:3}, w_{1:3}, \tau)\right) \\
\propto & \prod_{i=1}^{n} \left(\sum_{j=1}^{3} w_j N(Y_i; \mu_j, \tau)\right)
\end{aligned}
$$

Task 2

Derive the full conditionals for all the parameters up to a normalizing constant.

$$
\begin{aligned}
& p(w_{1:3} \mid Y_{1:n}, \mu_{1:3}, \mu_0, \phi_0, \tau) \\
\propto & \left(\prod_{i=1}^{n} p(Y_i \mid \mu_{1:3}, w_{1:3}, \tau) \right) \left(\prod_{j=1}^{3} p(\mu_j \mid \mu_0, \phi_0) \right) p(\mu_0) p(\phi_0) p(\tau) \\
\propto & \left(\prod_{i=1}^{n} p(Y_i \mid \mu_{1:3}, w_{1:3}, \tau) \right) \\
\propto & \prod_{i=1}^{n} \left(\sum_{j=1}^{3} w_j N(Y_i; \mu_j, \tau) \right)
\end{aligned}
$$

Task 2

Derive the full conditionals for all the parameters up to a normalizing constant.

$$
\begin{aligned}
& p(w_{1:3} \mid Y_{1:n}, \mu_{1:3}, \mu_0, \phi_0, \tau) \\
\propto & \left(\prod_{i=1}^{n} p(Y_i \mid \mu_{1:3}, w_{1:3}, \tau) \right) \left(\prod_{j=1}^{3} p(\mu_j \mid \mu_0, \phi_0) \right) p(\mu_0) p(\phi_0) p(\tau) \\
\propto & \left(\prod_{i=1}^{n} p(Y_i \mid \mu_{1:3}, w_{1:3}, \tau) \right) \\
\propto & \prod_{i=1}^{n} \left(\sum_{j=1}^{3} w_j N(Y_i; \mu_j, \tau) \right)
\end{aligned}
$$

Task 2

$$
\begin{aligned}
&p(\mu_j \mid Y_{1:n}, \mu_{(-j)}, \mu_0, \phi_0, w_{1:3}, \tau) \\
&\propto \left(\prod_{i=1}^{n} p(Y_i \mid \mu_{1:3}, w_{1:3}, \tau)\right) \left(\prod_{j=1}^{3} p(\mu_j \mid \mu_0, \phi_0)\right) p(\mu_0) p(\phi_0) p(\tau) \\
&\propto \left(\prod_{i=1}^{n} p(Y_i \mid \mu_{1:3}, w_{1:3}, \tau)\right) p(\mu_j \mid \mu_0, \phi_0) \\
&\propto \left(\prod_{i=1}^{n} \left(\sum_{j=1}^{3} w_j N(Y_i; \mu_j, \tau)\right)\right) p(\mu_j \mid \mu_0, \phi_0)
\end{aligned}
$$

Task 2

$$
\begin{aligned}
& p(\mu_j \mid Y_{1:n}, \mu_{(-j)}, \mu_0, \phi_0, w_{1:3}, \tau) \\
& \propto \left(\prod_{i=1}^{n} p(Y_i \mid \mu_{1:3}, w_{1:3}, \tau) \right) \left(\prod_{j=1}^{3} p(\mu_j \mid \mu_0, \phi_0) \right) p(\mu_0) p(\phi_0) p(\tau) \\
& \propto \left(\prod_{i=1}^{n} p(Y_i \mid \mu_{1:3}, w_{1:3}, \tau) \right) p(\mu_j \mid \mu_0, \phi_0) \\
& \propto \left(\prod_{i=1}^{n} \left(\sum_{j=1}^{3} w_j N(Y_i; \mu_j, \tau) \right) \right) p(\mu_j \mid \mu_0, \phi_0)
\end{aligned}
$$

Task 2

$$
\begin{aligned}
& p(\mu_j \mid Y_{1:n}, \mu_{(-j)}, \mu_0, \phi_0, w_{1:3}, \tau) \\
& \propto \left(\prod_{i=1}^{n} p(Y_i \mid \mu_{1:3}, w_{1:3}, \tau) \right) \left(\prod_{j=1}^{3} p(\mu_j \mid \mu_0, \phi_0) \right) p(\mu_0) p(\phi_0) p(\tau) \\
& \propto \left(\prod_{i=1}^{n} p(Y_i \mid \mu_{1:3}, w_{1:3}, \tau) \right) p(\mu_j \mid \mu_0, \phi_0) \\
& \propto \left(\prod_{i=1}^{n} \left(\sum_{j=1}^{3} w_j N(Y_i; \mu_j, \tau) \right) \right) p(\mu_j \mid \mu_0, \phi_0)
\end{aligned}
$$

Task 2

$$
\begin{aligned}
& p(\mu_j \mid Y_{1:n}, \mu_{(-j)}, \mu_0, \phi_0, w_{1:3}, \tau) \\
\propto & \left(\prod_{i=1}^{n} p(Y_i \mid \mu_{1:3}, w_{1:3}, \tau) \right) \left(\prod_{j=1}^{3} p(\mu_j \mid \mu_0, \phi_0) \right) p(\mu_0) p(\phi_0) p(\tau) \\
\propto & \left(\prod_{i=1}^{n} p(Y_i \mid \mu_{1:3}, w_{1:3}, \tau) \right) p(\mu_j \mid \mu_0, \phi_0) \\
\propto & \left(\prod_{i=1}^{n} \left(\sum_{j=1}^{3} w_j N(Y_i; \mu_j, \tau) \right) \right) p(\mu_j \mid \mu_0, \phi_0)
\end{aligned}
$$

Task 2

$$
\begin{aligned}
& p(\tau \mid Y_{1:n}, \mu_{1:3}, \mu_0, \phi_0, w_{1:3}) \\
& \propto \left(\prod_{i=1}^{n} p(Y_i \mid \mu_{1:3}, w_{1:3}, \tau)\right) \left(\prod_{j=1}^{3} p(\mu_j \mid \mu_0, \phi_0)\right) p(\mu_0) p(\phi_0) p(\tau) \\
& \propto \left(\prod_{i=1}^{n} p(Y_i \mid \mu_{1:3}, w_{1:3}, \tau)\right) p(\tau) \\
& \propto \left(\prod_{i=1}^{n} \left(\sum_{j=1}^{3} w_j N(Y_i; \mu_j, \tau)\right)\right) \tau^{2-1} \exp\{-2\tau\}.
\end{aligned}
$$

Task 2

$$
\begin{aligned}
& p(\tau \mid Y_{1:n}, \mu_{1:3}, \mu_0, \phi_0, w_{1:3}) \\
\propto & \left(\prod_{i=1}^{n} p(Y_i \mid \mu_{1:3}, w_{1:3}, \tau) \right) \left(\prod_{j=1}^{3} p(\mu_j \mid \mu_0, \phi_0) \right) p(\mu_0) p(\phi_0) p(\tau) \\
\propto & \left(\prod_{i=1}^{n} p(Y_i \mid \mu_{1:3}, w_{1:3}, \tau) \right) p(\tau) \\
\propto & \left(\prod_{i=1}^{n} \left(\sum_{j=1}^{3} w_j N(Y_i; \mu_j, \tau) \right) \right) \tau^{2-1} \exp\{-2\tau\}.
\end{aligned}
$$

Task 2

$$
\begin{aligned}
&p(\tau \mid Y_{1:n}, \mu_{1:3}, \mu_0, \phi_0, w_{1:3}) \\
&\propto \left(\prod_{i=1}^{n} p(Y_i \mid \mu_{1:3}, w_{1:3}, \tau) \right) \left(\prod_{j=1}^{3} p(\mu_j \mid \mu_0, \phi_0) \right) p(\mu_0) p(\phi_0) p(\tau) \\
&\propto \left(\prod_{i=1}^{n} p(Y_i \mid \mu_{1:3}, w_{1:3}, \tau) \right) p(\tau) \\
&\propto \left(\prod_{i=1}^{n} \left(\sum_{j=1}^{3} w_j N(Y_i; \mu_j, \tau) \right) \right) \tau^{2-1} \exp\{-2\tau\}.
\end{aligned}
$$

Task 2

$$
\begin{aligned}
& p(\tau \mid Y_{1:n}, \mu_{1:3}, \mu_0, \phi_0, w_{1:3}) \\
\propto & \left(\prod_{i=1}^{n} p(Y_i \mid \mu_{1:3}, w_{1:3}, \tau) \right) \left(\prod_{j=1}^{3} p(\mu_j \mid \mu_0, \phi_0) \right) p(\mu_0) p(\phi_0) p(\tau) \\
\propto & \left(\prod_{i=1}^{n} p(Y_i \mid \mu_{1:3}, w_{1:3}, \tau) \right) p(\tau) \\
\propto & \left(\prod_{i=1}^{n} \left(\sum_{j=1}^{3} w_j N(Y_i; \mu_j, \tau) \right) \right) \tau^{2-1} \exp\{-2\tau\}.
\end{aligned}
$$

Task 2

$$
\begin{aligned}
& p(\mu_0 \mid Y_{1:n}, \mu_{1:3}, \phi_0, w_{1:3}, \tau) \\
\propto & \left(\prod_{i=1}^{n} p(Y_i \mid \mu_{1:3}, w_{1:3}, \tau) \right) \left(\prod_{j=1}^{3} p(\mu_j \mid \mu_0, \phi_0) \right) p(\mu_0) p(\phi_0) p(\tau) \\
\propto & \left(\prod_{j=1}^{3} p(\mu_j \mid \mu_0, \phi_0) \right) p(\mu_0) \\
\propto & \exp\left\{ -\frac{\phi_0}{2} \sum_{j=1}^{3} (\mu_j - \mu_0)^2 \right\} \exp\{\mu_0^2/6\} \\
\propto & \exp\left\{ -\frac{1}{2} \left[(3\phi_0 + \frac{1}{3})\mu_0^2 - 2\mu_0\phi_0 \sum_{j=1}^{3} \mu_j \right] \right\} \\
\Rightarrow & \mu_0 \mid - \sim N\left((3\phi_0 + \frac{1}{3})^{-1} \phi_0 \sum_{j=1}^{3} \mu_j, (3\phi_0 + \frac{1}{3})^{-1} \right).
\end{aligned}
$$

Task 2

$$
\begin{aligned}
& p(\mu_0 \mid Y_{1:n}, \mu_{1:3}, \phi_0, w_{1:3}, \tau) \\
\propto & \left(\prod_{i=1}^{n} p(Y_i \mid \mu_{1:3}, w_{1:3}, \tau) \right) \left(\prod_{j=1}^{3} p(\mu_j \mid \mu_0, \phi_0) \right) p(\mu_0) p(\phi_0) p(\tau) \\
\propto & \left(\prod_{j=1}^{3} p(\mu_j \mid \mu_0, \phi_0) \right) p(\mu_0) \\
\propto & \exp\left\{ -\frac{\phi_0}{2} \sum_{j=1}^{3} (\mu_j - \mu_0)^2 \right\} \exp\{\mu_0^2/6\} \\
\propto & \exp\left\{ -\frac{1}{2} \left[(3\phi_0 + \frac{1}{3})\mu_0^2 - 2\mu_0\phi_0 \sum_{j=1}^{3} \mu_j \right] \right\} \\
\Rightarrow & \mu_0 \mid - \sim N\left((3\phi_0 + \frac{1}{3})^{-1} \phi_0 \sum_{j=1}^{3} \mu_j, (3\phi_0 + \frac{1}{3})^{-1} \right).
\end{aligned}
$$

Task 2

$$
\begin{aligned}
& p(\mu_0 \mid Y_{1:n}, \mu_{1:3}, \phi_0, w_{1:3}, \tau) \\
\propto & \left(\prod_{i=1}^{n} p(Y_i \mid \mu_{1:3}, w_{1:3}, \tau)\right) \left(\prod_{j=1}^{3} p(\mu_j \mid \mu_0, \phi_0)\right) p(\mu_0)p(\phi_0)p(\tau) \\
\propto & \left(\prod_{j=1}^{3} p(\mu_j \mid \mu_0, \phi_0)\right) p(\mu_0) \\
\propto & \exp\left\{-\frac{\phi_0}{2}\sum_{j=1}^{3}(\mu_j - \mu_0)^2\right\} \exp\{\mu_0^2/6\} \\
\propto & \exp\left\{-\frac{1}{2}\left[(3\phi_0 + \frac{1}{3})\mu_0^2 - 2\mu_0\phi_0\sum_{j=1}^{3}\mu_j\right]\right\} \\
\Rightarrow & \mu_0 \mid - \sim N\left((3\phi_0 + \frac{1}{3})^{-1}\phi_0\sum_{j=1}^{3}\mu_j, (3\phi_0 + \frac{1}{3})^{-1}\right).
\end{aligned}
$$

Task 2

$$
\begin{aligned}
& p(\mu_0 \mid Y_{1:n}, \mu_{1:3}, \phi_0, w_{1:3}, \tau) \\
\propto & \left(\prod_{i=1}^{n} p(Y_i \mid \mu_{1:3}, w_{1:3}, \tau) \right) \left(\prod_{j=1}^{3} p(\mu_j \mid \mu_0, \phi_0) \right) p(\mu_0) p(\phi_0) p(\tau) \\
\propto & \left(\prod_{j=1}^{3} p(\mu_j \mid \mu_0, \phi_0) \right) p(\mu_0) \\
\propto & \exp \left\{ -\frac{\phi_0}{2} \sum_{j=1}^{3} (\mu_j - \mu_0)^2 \right\} \exp\{\mu_0^2/6\} \\
\propto & \exp \left\{ -\frac{1}{2} \left[(3\phi_0 + \frac{1}{3})\mu_0^2 - 2\mu_0\phi_0 \sum_{j=1}^{3} \mu_j \right] \right\} \\
\Rightarrow & \mu_0 \mid - \sim N \left((3\phi_0 + \frac{1}{3})^{-1} \phi_0 \sum_{j=1}^{3} \mu_j, (3\phi_0 + \frac{1}{3})^{-1} \right).
\end{aligned}
$$

Task 2

$$
\begin{aligned}
& p(\mu_0 \mid Y_{1:n}, \mu_{1:3}, \phi_0, w_{1:3}, \tau) \\
\propto & \left(\prod_{i=1}^{n} p(Y_i \mid \mu_{1:3}, w_{1:3}, \tau)\right) \left(\prod_{j=1}^{3} p(\mu_j \mid \mu_0, \phi_0)\right) p(\mu_0) p(\phi_0) p(\tau) \\
\propto & \left(\prod_{j=1}^{3} p(\mu_j \mid \mu_0, \phi_0)\right) p(\mu_0) \\
\propto & \exp\left\{-\frac{\phi_0}{2}\sum_{j=1}^{3}(\mu_j - \mu_0)^2\right\} \exp\{\mu_0^2/6\} \\
\propto & \exp\left\{-\frac{1}{2}\left[(3\phi_0 + \frac{1}{3})\mu_0^2 - 2\mu_0\phi_0\sum_{j=1}^{3}\mu_j\right]\right\} \\
\Rightarrow & \mu_0 \mid - \sim N\left((3\phi_0 + \frac{1}{3})^{-1}\phi_0\sum_{j=1}^{3}\mu_j, (3\phi_0 + \frac{1}{3})^{-1}\right).
\end{aligned}
$$

Task 2

$$
\begin{aligned}
& p(\mu_0 \mid Y_{1:n}, \mu_{1:3}, \phi_0, w_{1:3}, \tau) \\
\propto & \left(\prod_{i=1}^{n} p(Y_i \mid \mu_{1:3}, w_{1:3}, \tau) \right) \left(\prod_{j=1}^{3} p(\mu_j \mid \mu_0, \phi_0) \right) p(\mu_0) p(\phi_0) p(\tau) \\
\propto & \left(\prod_{j=1}^{3} p(\mu_j \mid \mu_0, \phi_0) \right) p(\mu_0) \\
\propto & \exp\left\{ -\frac{\phi_0}{2} \sum_{j=1}^{3} (\mu_j - \mu_0)^2 \right\} \exp\{\mu_0^2/6\} \\
\propto & \exp\left\{ -\frac{1}{2} \left[(3\phi_0 + \frac{1}{3})\mu_0^2 - 2\mu_0\phi_0 \sum_{j=1}^{3} \mu_j \right] \right\} \\
\Rightarrow & \mu_0 \mid - \sim N\left((3\phi_0 + \frac{1}{3})^{-1} \phi_0 \sum_{j=1}^{3} \mu_j, (3\phi_0 + \frac{1}{3})^{-1} \right).
\end{aligned}
$$

Task 2

$$
\begin{aligned}
& p(\phi_0 \mid Y_{1:n}, \mu_{1:3}, \mu_0, w_{1:3}, \tau) \\
\propto & \left(\prod_{i=1}^{n} p(Y_i \mid \mu_{1:3}, w_{1:3}, \tau) \right) \left(\prod_{j=1}^{3} p(\mu_j \mid \mu_0, \phi_0) \right) p(\mu_0) p(\phi_0) p(\tau) \\
\propto & \left(\prod_{j=1}^{3} p(\mu_j \mid \mu_0, \phi_0) \right) p(\phi_0) \\
\propto & \left(\prod_{j=1}^{3} \sqrt{\frac{\phi_0}{2\pi}} \exp\left\{ -\frac{\phi_0}{2} (\mu_0 - \mu_j)^2 \right\} \right) \phi_0^{2-1} \exp\{-2\phi_0\} \\
\propto & \phi_0^{7/2-1} \exp\left\{ -\phi_0 \left(2 + \frac{1}{2} \sum_{j=1}^{3} (\mu_0 - \mu_j)^2 \right) \right\} \\
\Rightarrow & \phi_0 \mid - \sim \text{Gamma}\left(7/2, 2 + \frac{1}{2} \sum_{j=1}^{3} (\mu_0 - \mu_j)^2 \right)
\end{aligned}
$$

Task 2

$$
\begin{aligned}
& p(\phi_0 \mid Y_{1:n}, \mu_{1:3}, \mu_0, w_{1:3}, \tau) \\
\propto & \left(\prod_{i=1}^{n} p(Y_i \mid \mu_{1:3}, w_{1:3}, \tau)\right) \left(\prod_{j=1}^{3} p(\mu_j \mid \mu_0, \phi_0)\right) p(\mu_0) p(\phi_0) p(\tau) \\
\propto & \left(\prod_{j=1}^{3} p(\mu_j \mid \mu_0, \phi_0)\right) p(\phi_0) \\
\propto & \left(\prod_{j=1}^{3} \sqrt{\frac{\phi_0}{2\pi}} \exp\left\{-\frac{\phi_0}{2}(\mu_0 - \mu_j)^2\right\}\right) \phi_0^{2-1} \exp\{-2\phi_0\} \\
\propto & \phi_0^{7/2-1} \exp\left\{-\phi_0 \left(2 + \frac{1}{2}\sum_{j=1}^{3}(\mu_0 - \mu_j)^2\right)\right\} \\
\Rightarrow & \phi_0 \mid - \sim \text{Gamma}\left(7/2, 2 + \frac{1}{2}\sum_{j=1}^{3}(\mu_0 - \mu_j)^2\right)
\end{aligned}
$$

Task 2

$$
\begin{aligned}
& p(\phi_0 \mid Y_{1:n}, \mu_{1:3}, \mu_0, w_{1:3}, \tau) \\
\propto & \left(\prod_{i=1}^{n} p(Y_i \mid \mu_{1:3}, w_{1:3}, \tau) \right) \left(\prod_{j=1}^{3} p(\mu_j \mid \mu_0, \phi_0) \right) p(\mu_0) p(\phi_0) p(\tau) \\
\propto & \left(\prod_{j=1}^{3} p(\mu_j \mid \mu_0, \phi_0) \right) p(\phi_0) \\
\propto & \left(\prod_{j=1}^{3} \sqrt{\frac{\phi_0}{2\pi}} \exp\left\{ -\frac{\phi_0}{2} (\mu_0 - \mu_j)^2 \right\} \right) \phi_0^{2-1} \exp\{-2\phi_0\} \\
\propto & \phi_0^{7/2-1} \exp\left\{ -\phi_0 \left(2 + \frac{1}{2} \sum_{j=1}^{3} (\mu_0 - \mu_j)^2 \right) \right\} \\
\Rightarrow & \phi_0 \mid - \sim \text{Gamma}\left(7/2, 2 + \frac{1}{2} \sum_{j=1}^{3} (\mu_0 - \mu_j)^2 \right)
\end{aligned}
$$

Task 2

$$
\begin{aligned}
& p(\phi_0 \mid Y_{1:n}, \mu_{1:3}, \mu_0, w_{1:3}, \tau) \\
\propto & \left(\prod_{i=1}^{n} p(Y_i \mid \mu_{1:3}, w_{1:3}, \tau)\right) \left(\prod_{j=1}^{3} p(\mu_j \mid \mu_0, \phi_0)\right) p(\mu_0) p(\phi_0) p(\tau) \\
\propto & \left(\prod_{j=1}^{3} p(\mu_j \mid \mu_0, \phi_0)\right) p(\phi_0) \\
\propto & \left(\prod_{j=1}^{3} \sqrt{\frac{\phi_0}{2\pi}} \exp\left\{-\frac{\phi_0}{2}(\mu_0 - \mu_j)^2\right\}\right) \phi_0^{2-1} \exp\{-2\phi_0\} \\
\propto & \phi_0^{7/2-1} \exp\left\{-\phi_0 \left(2 + \frac{1}{2}\sum_{j=1}^{3}(\mu_0 - \mu_j)^2\right)\right\} \\
\Rightarrow & \phi_0 \mid - \sim \text{Gamma}\left(7/2, 2 + \frac{1}{2}\sum_{j=1}^{3}(\mu_0 - \mu_j)^2\right)
\end{aligned}
$$

Task 2

$$
\begin{aligned}
& p(\phi_0 \mid Y_{1:n}, \mu_{1:3}, \mu_0, w_{1:3}, \tau) \\
\propto & \left(\prod_{i=1}^{n} p(Y_i \mid \mu_{1:3}, w_{1:3}, \tau) \right) \left(\prod_{j=1}^{3} p(\mu_j \mid \mu_0, \phi_0) \right) p(\mu_0) p(\phi_0) p(\tau) \\
\propto & \left(\prod_{j=1}^{3} p(\mu_j \mid \mu_0, \phi_0) \right) p(\phi_0) \\
\propto & \left(\prod_{j=1}^{3} \sqrt{\frac{\phi_0}{2\pi}} \exp\left\{ -\frac{\phi_0}{2} (\mu_0 - \mu_j)^2 \right\} \right) \phi_0^{2-1} \exp\{-2\phi_0\} \\
\propto & \phi_0^{7/2-1} \exp\left\{ -\phi_0 \left(2 + \frac{1}{2} \sum_{j=1}^{3} (\mu_0 - \mu_j)^2 \right) \right\} \\
\Rightarrow & \phi_0 \mid - \sim \text{Gamma}\left(7/2, 2 + \frac{1}{2} \sum_{j=1}^{3} (\mu_0 - \mu_j)^2 \right)
\end{aligned}
$$

Task 2

$$
\begin{aligned}
& p(\phi_0 \mid Y_{1:n}, \mu_{1:3}, \mu_0, w_{1:3}, \tau) \\
\propto & \left(\prod_{i=1}^{n} p(Y_i \mid \mu_{1:3}, w_{1:3}, \tau)\right) \left(\prod_{j=1}^{3} p(\mu_j \mid \mu_0, \phi_0)\right) p(\mu_0) p(\phi_0) p(\tau) \\
\propto & \left(\prod_{j=1}^{3} p(\mu_j \mid \mu_0, \phi_0)\right) p(\phi_0) \\
\propto & \left(\prod_{j=1}^{3} \sqrt{\frac{\phi_0}{2\pi}} \exp\left\{-\frac{\phi_0}{2}(\mu_0 - \mu_j)^2\right\}\right) \phi_0^{2-1} \exp\{-2\phi_0\} \\
\propto & \phi_0^{7/2-1} \exp\left\{-\phi_0 \left(2 + \frac{1}{2}\sum_{j=1}^{3}(\mu_0 - \mu_j)^2\right)\right\} \\
\Rightarrow & \phi_0 \mid - \sim \text{Gamma}\left(7/2, 2 + \frac{1}{2}\sum_{j=1}^{3}(\mu_0 - \mu_j)^2\right)
\end{aligned}
$$

Task 3

Where necessary, (re)-derive the full conditionals under the data augmentation scheme.

Data augmentation scheme:

- Same priors as before, but we decompose the likelihood:

$$Z_i \mid w_1, w_2, w_3 \sim Cat(3, \boldsymbol{w})$$
$$Y_i \mid Z_i, \mu_{1:3}, \tau \sim N(\mu_{Z_i}, 1/\tau)$$

- The marginal distribution of Y_i is **unchanged**.
- Using the variables Z_i helps derive full conditional distributions.

In other words

$$p(Y_i \mid Z_i, \mu_1, \mu_2, \mu_3, \varepsilon^2) = N(Y_i; \mu_{Z_i}, \varepsilon^2)$$
$$P(Z_i = j \mid w_{1:3}) = w_j.$$

Task 3

Where necessary, (re)-derive the full conditionals under the data augmentation scheme.

Data augmentation scheme:

- Same priors as before, but we decompose the likelihood:

$$Z_i \mid w_1, w_2, w_3 \sim Cat(3, \boldsymbol{w})$$
$$Y_i \mid Z_i, \mu_{1:3}, \tau \sim N(\mu_{Z_i}, 1/\tau)$$

- The marginal distribution of Y_i is **unchanged**.
- Using the variables Z_i helps derive full conditional distributions.

In other words

$$p(Y_i \mid Z_i, \mu_1, \mu_2, \mu_3, \varepsilon^2) = N(Y_i; \mu_{Z_i}, \varepsilon^2)$$
$$P(Z_i = j \mid w_{1:3}) = w_j.$$

Task 3

Where necessary, (re)-derive the full conditionals under the data augmentation scheme.

Data augmentation scheme:

- Same priors as before, but we decompose the likelihood:

$$Z_i \mid w_1, w_2, w_3 \sim Cat(3, \boldsymbol{w})$$
$$Y_i \mid Z_i, \mu_{1:3}, \tau \sim N(\mu_{Z_i}, 1/\tau)$$

- The marginal distribution of Y_i is **unchanged**.
- Using the variables Z_i helps derive full conditional distributions.

In other words

$$p(Y_i \mid Z_i, \mu_1, \mu_2, \mu_3, \varepsilon^2) = N(Y_i; \mu_{Z_i}, \varepsilon^2)$$
$$P(Z_i = j \mid w_{1:3}) = w_j.$$

Task 3

Where necessary, (re)-derive the full conditionals under the data augmentation scheme.

Data augmentation scheme:

- Same priors as before, but we decompose the likelihood:

$$Z_i \mid w_1, w_2, w_3 \sim Cat(3, \boldsymbol{w})$$
$$Y_i \mid Z_i, \mu_{1:3}, \tau \sim N(\mu_{Z_i}, 1/\tau)$$

- The marginal distribution of Y_i is **unchanged**.
- Using the variables Z_i helps derive full conditional distributions.

In other words

$$p(Y_i \mid Z_i, \mu_1, \mu_2, \mu_3, \varepsilon^2) = N(Y_i; \mu_{Z_i}, \varepsilon^2)$$
$$P(Z_i = j \mid w_{1:3}) = w_j.$$

Task 3

Where necessary, (re)-derive the full conditionals under the data augmentation scheme.

Data augmentation scheme:

- Same priors as before, but we decompose the likelihood:

$$Z_i \mid w_1, w_2, w_3 \sim Cat(3, \boldsymbol{w})$$
$$Y_i \mid Z_i, \mu_{1:3}, \tau \sim N(\mu_{Z_i}, 1/\tau)$$

- The marginal distribution of Y_i is **unchanged**.
- Using the variables Z_i helps derive full conditional distributions.

In other words

$$p(Y_i \mid Z_i, \mu_1, \mu_2, \mu_3, \varepsilon^2) = N(Y_i; \mu_{Z_i}, \varepsilon^2)$$
$$P(Z_i = j \mid w_{1:3}) = w_j.$$

Task 3

Where necessary, (re)-derive the full conditionals under the data augmentation scheme.

Data augmentation scheme:

- Same priors as before, but we decompose the likelihood:

$$
\begin{aligned}
Z_i \mid w_1, w_2, w_3 &\sim Cat(3, \boldsymbol{w}) \\
Y_i \mid Z_i, \mu_{1:3}, \tau &\sim N(\mu_{Z_i}, 1/\tau)
\end{aligned}
$$

- The marginal distribution of Y_i is **unchanged**.
- Using the variables Z_i helps derive full conditional distributions.

In other words

$$
\begin{aligned}
p(Y_i \mid Z_i, \mu_1, \mu_2, \mu_3, \varepsilon^2) &= N(Y_i; \mu_{Z_i}, \varepsilon^2) \\
P(Z_i = j \mid w_{1:3}) &= w_j.
\end{aligned}
$$

Task 3

Where necessary, (re)-derive the full conditionals under the data augmentation scheme.

Data augmentation scheme:

- Same priors as before, but we decompose the likelihood:

$$Z_i \mid w_1, w_2, w_3 \sim Cat(3, \boldsymbol{w})$$
$$Y_i \mid Z_i, \mu_{1:3}, \tau \sim N(\mu_{Z_i}, 1/\tau)$$

- The marginal distribution of Y_i is **unchanged**.
- Using the variables Z_i helps derive full conditional distributions.

In other words

$$p(Y_i \mid Z_i, \mu_1, \mu_2, \mu_3, \varepsilon^2) = N(Y_i; \mu_{Z_i}, \varepsilon^2)$$
$$P(Z_i = j \mid w_{1:3}) = w_j.$$

Task 3

Now the full joint becomes:

$$
\begin{aligned}
&p(Y_{1:n}, Z_{1:n}, \mu_{1:3}, \mu_0, \phi_0, w_{1:3}, \tau) \\
=&p(Y_{1:n} \mid Z_{1:n}, \mu_{1:3}, \tau) p(Z_{1:n}) p(\mu_{1:3} \mid \mu_0, \phi_0) p(\mu_0) p(\phi_0) p(\tau) \\
=&\prod_{i=1}^{n} \left(N(Y_i; \mu_{Z_i}, \tau^{-1}) w_{Z_i} \right) p(\mu_{1:3} \mid \mu_0, \phi_0) p(\mu_0) p(\phi_0) p(\tau)
\end{aligned}
$$

Task 3

Now the full joint becomes:

$$
\begin{aligned}
&p(Y_{1:n}, Z_{1:n}, \mu_{1:3}, \mu_0, \phi_0, w_{1:3}, \tau) \\
=&p(Y_{1:n} \mid Z_{1:n}, \mu_{1:3}, \tau) p(Z_{1:n}) p(\mu_{1:3} \mid \mu_0, \phi_0) p(\mu_0) p(\phi_0) p(\tau) \\
=&\prod_{i=1}^{n} \left(N(Y_i; \mu_{Z_i}, \tau^{-1}) w_{Z_i} \right) p(\mu_{1:3} \mid \mu_0, \phi_0) p(\mu_0) p(\phi_0) p(\tau)
\end{aligned}
$$

Task 3

Now the full joint becomes:

$$\begin{aligned}
&p(Y_{1:n}, Z_{1:n}, \mu_{1:3}, \mu_0, \phi_0, w_{1:3}, \tau) \\
=&p(Y_{1:n} \mid Z_{1:n}, \mu_{1:3}, \tau)p(Z_{1:n})p(\mu_{1:3} \mid \mu_0, \phi_0)p(\mu_0)p(\phi_0)p(\tau) \\
=&\prod_{i=1}^{n} \left(N(Y_i; \mu_{Z_i}, \tau^{-1})w_{Z_i}\right) p(\mu_{1:3} \mid \mu_0, \phi_0)p(\mu_0)p(\phi_0)p(\tau)
\end{aligned}$$

Task 3

Full conditional of $w_{1:3}$:

$$
\begin{aligned}
& p(w_{1:3} \mid Z_{1:n}, Y_{1:n}, \mu_{1:3}, \mu_0, \phi_0, \tau) \\
& \propto \prod_{i=1}^{n} \left(N(Y_i; \mu_{Z_i}, \tau^{-1}) w_{Z_i} \right) p(\mu_{1:3} \mid \mu_0, \phi_0) p(\mu_0) p(\phi_0) p(\tau) \\
& \propto \prod_{i=1}^{n} w_{Z_i} \\
& \propto w_1^{N_1} w_2^{N_2} w_3^{N_3} \\
& \Rightarrow w_{1:3} \sim \text{Dirichlet}(N_1 + 1, N_2 + 1, N_3 + 1)
\end{aligned}
$$

where $N_j = \sum_{i=1}^{n} \mathbb{I}(Z_i = j)$.

Task 3

Full conditional of $w_{1:3}$:

$$
\begin{aligned}
& p(w_{1:3} \mid Z_{1:n}, Y_{1:n}, \mu_{1:3}, \mu_0, \phi_0, \tau) \\
\propto & \prod_{i=1}^{n} \left(N(Y_i; \mu_{Z_i}, \tau^{-1}) w_{Z_i} \right) p(\mu_{1:3} \mid \mu_0, \phi_0) p(\mu_0) p(\phi_0) p(\tau) \\
\propto & \prod_{i=1}^{n} w_{Z_i} \\
\propto & w_1^{N_1} w_2^{N_2} w_3^{N_3} \\
\Rightarrow & w_{1:3} \sim \text{Dirichlet}(N_1 + 1, N_2 + 1, N_3 + 1)
\end{aligned}
$$

where $N_j = \sum_{i=1}^{n} \mathbb{I}(Z_i = j)$.

Task 3

Full conditional of $w_{1:3}$:

$$
\begin{aligned}
& p(w_{1:3} \mid Z_{1:n}, Y_{1:n}, \mu_{1:3}, \mu_0, \phi_0, \tau) \\
\propto & \prod_{i=1}^{n} \left(N(Y_i; \mu_{Z_i}, \tau^{-1}) w_{Z_i} \right) p(\mu_{1:3} \mid \mu_0, \phi_0) p(\mu_0) p(\phi_0) p(\tau) \\
\propto & \prod_{i=1}^{n} w_{Z_i} \\
\propto & \, w_1^{N_1} w_2^{N_2} w_3^{N_3} \\
\Rightarrow & \, w_{1:3} \sim \text{Dirichlet}(N_1 + 1, N_2 + 1, N_3 + 1)
\end{aligned}
$$

where $N_j = \sum_{i=1}^{n} \mathbb{I}(Z_i = j)$.

Task 3

Full conditional of $w_{1:3}$:

$$
\begin{aligned}
& p(w_{1:3} \mid Z_{1:n}, Y_{1:n}, \mu_{1:3}, \mu_0, \phi_0, \tau) \\
\propto & \prod_{i=1}^{n} \left(N(Y_i; \mu_{Z_i}, \tau^{-1}) w_{Z_i} \right) p(\mu_{1:3} \mid \mu_0, \phi_0) p(\mu_0) p(\phi_0) p(\tau) \\
\propto & \prod_{i=1}^{n} w_{Z_i} \\
\propto & w_1^{N_1} w_2^{N_2} w_3^{N_3} \\
\Rightarrow & w_{1:3} \sim \text{Dirichlet}(N_1 + 1, N_2 + 1, N_3 + 1)
\end{aligned}
$$

where $N_j = \sum_{i=1}^{n} \mathbb{I}(Z_i = j)$.

Task 3

Full conditional of $w_{1:3}$:

$$
\begin{aligned}
& p(w_{1:3} \mid Z_{1:n}, Y_{1:n}, \mu_{1:3}, \mu_0, \phi_0, \tau) \\
\propto & \prod_{i=1}^{n} \left(N(Y_i; \mu_{Z_i}, \tau^{-1}) w_{Z_i} \right) p(\mu_{1:3} \mid \mu_0, \phi_0) p(\mu_0) p(\phi_0) p(\tau) \\
\propto & \prod_{i=1}^{n} w_{Z_i} \\
\propto & w_1^{N_1} w_2^{N_2} w_3^{N_3} \\
\Rightarrow & w_{1:3} \sim \text{Dirichlet}(N_1 + 1, N_2 + 1, N_3 + 1)
\end{aligned}
$$

where $N_j = \sum_{i=1}^{n} \mathbb{I}(Z_i = j)$.

Task 3

Full conditional of μ_j:

$$
\begin{aligned}
& p(\mu_j \mid Y_{1:n}, Z_{1:n}, \mu_{(-j)}, \mu_0, \phi_0, w_{1:3}, \tau) \\
& \propto \prod_{i=1}^{n} \left(N(Y_i; \mu_{Z_i}, \tau^{-1}) w_{Z_i} \right) p(\mu_{1:3} \mid \mu_0, \phi_0) p(\mu_0) p(\phi_0) p(\tau) \\
& \propto \prod_{i=1}^{n} \left(N(Y_i; \mu_{Z_i}, \tau^{-1}) \right) p(\mu_j \mid \mu_0, \phi_0) \\
& \propto \exp\left\{ \frac{-\tau}{2} \sum_{i:Z_i=j} (Y_i - \mu_j)^2 \right\} \exp\{ \frac{-\phi_0}{2} (\mu_j - \mu_0)^2 \} \\
& \propto \exp\left\{ \frac{-1}{2} \left[\mu_j^2 (\tau N_j + \phi_0) - 2\mu_j (\tau \sum_{i:Z_i=j} Y_i + \phi_0 \mu_0) \mu_j \right] \right\} \\
& \Rightarrow \mu_j \mid - \sim N((\tau N_j + \phi_0)^{-1} (\tau \sum_{i:Z_i=j} Y_i + \phi_0 \mu_0), (\tau N_j + \phi_0)^{-1})
\end{aligned}
$$

Task 3

Full conditional of μ_j:

$$
\begin{aligned}
& p(\mu_j \mid Y_{1:n}, Z_{1:n}, \mu_{(-j)}, \mu_0, \phi_0, w_{1:3}, \tau) \\
\propto & \prod_{i=1}^{n} \left(N(Y_i; \mu_{Z_i}, \tau^{-1}) w_{Z_i}\right) p(\mu_{1:3} \mid \mu_0, \phi_0) p(\mu_0) p(\phi_0) p(\tau) \\
\propto & \prod_{i=1}^{n} \left(N(Y_i; \mu_{Z_i}, \tau^{-1})\right) p(\mu_j \mid \mu_0, \phi_0) \\
\propto & \exp\left\{ \frac{-\tau}{2} \sum_{i:Z_i=j} (Y_i - \mu_j)^2 \right\} \exp\{ \frac{-\phi_0}{2} (\mu_j - \mu_0)^2 \} \\
\propto & \exp\left\{ \frac{-1}{2} \left[\mu_j^2 (\tau N_j + \phi_0) - 2\mu_j (\tau \sum_{i:Z_i=j} Y_i + \phi_0 \mu_0) \mu_j \right] \right\} \\
\Rightarrow & \mu_j \mid - \sim N((\tau N_j + \phi_0)^{-1} (\tau \sum_{i:Z_i=j} Y_i + \phi_0 \mu_0), (\tau N_j + \phi_0)^{-1})
\end{aligned}
$$

Task 3

Full conditional of μ_j:

$$
\begin{aligned}
& p(\mu_j \mid Y_{1:n}, Z_{1:n}, \mu_{(-j)}, \mu_0, \phi_0, w_{1:3}, \tau) \\
\propto & \prod_{i=1}^{n} \left(N(Y_i; \mu_{Z_i}, \tau^{-1}) w_{Z_i} \right) p(\mu_{1:3} \mid \mu_0, \phi_0) p(\mu_0) p(\phi_0) p(\tau) \\
\propto & \prod_{i=1}^{n} \left(N(Y_i; \mu_{Z_i}, \tau^{-1}) \right) p(\mu_j \mid \mu_0, \phi_0) \\
\propto & \exp\left\{ \frac{-\tau}{2} \sum_{i:Z_i=j} (Y_i - \mu_j)^2 \right\} \exp\{ \frac{-\phi_0}{2} (\mu_j - \mu_0)^2 \} \\
\propto & \exp\left\{ \frac{-1}{2} \left[\mu_j^2 (\tau N_j + \phi_0) - 2\mu_j (\tau \sum_{i:Z_i=j} Y_i + \phi_0 \mu_0) \mu_j \right] \right\} \\
\Rightarrow & \mu_j \mid - \sim N((\tau N_j + \phi_0)^{-1} (\tau \sum_{i:Z_i=j} Y_i + \phi_0 \mu_0), (\tau N_j + \phi_0)^{-1})
\end{aligned}
$$

Task 3

Full conditional of μ_j:

$$
\begin{aligned}
& p(\mu_j \mid Y_{1:n}, Z_{1:n}, \mu_{(-j)}, \mu_0, \phi_0, w_{1:3}, \tau) \\
\propto & \prod_{i=1}^{n} \left(N(Y_i; \mu_{Z_i}, \tau^{-1}) w_{Z_i} \right) p(\mu_{1:3} \mid \mu_0, \phi_0) p(\mu_0) p(\phi_0) p(\tau) \\
\propto & \prod_{i=1}^{n} \left(N(Y_i; \mu_{Z_i}, \tau^{-1}) \right) p(\mu_j \mid \mu_0, \phi_0) \\
\propto & \exp\left\{ \frac{-\tau}{2} \sum_{i:Z_i=j} (Y_i - \mu_j)^2 \right\} \exp\{ \frac{-\phi_0}{2} (\mu_j - \mu_0)^2 \} \\
\propto & \exp\left\{ \frac{-1}{2} \left[\mu_j^2 (\tau N_j + \phi_0) - 2\mu_j (\tau \sum_{i:Z_i=j} Y_i + \phi_0 \mu_0) \mu_j \right] \right\} \\
\Rightarrow & \mu_j \mid - \sim N((\tau N_j + \phi_0)^{-1} (\tau \sum_{i:Z_i=j} Y_i + \phi_0 \mu_0), (\tau N_j + \phi_0)^{-1})
\end{aligned}
$$

Task 3

Full conditional of μ_j:

$$
\begin{aligned}
& p(\mu_j \mid Y_{1:n}, Z_{1:n}, \mu_{(-j)}, \mu_0, \phi_0, w_{1:3}, \tau) \\
& \propto \prod_{i=1}^{n} \left(N(Y_i; \mu_{Z_i}, \tau^{-1}) w_{Z_i} \right) p(\mu_{1:3} \mid \mu_0, \phi_0) p(\mu_0) p(\phi_0) p(\tau) \\
& \propto \prod_{i=1}^{n} \left(N(Y_i; \mu_{Z_i}, \tau^{-1}) \right) p(\mu_j \mid \mu_0, \phi_0) \\
& \propto \exp\left\{ \frac{-\tau}{2} \sum_{i:Z_i=j} (Y_i - \mu_j)^2 \right\} \exp\{ \frac{-\phi_0}{2} (\mu_j - \mu_0)^2 \} \\
& \propto \exp\left\{ \frac{-1}{2} \left[\mu_j^2 (\tau N_j + \phi_0) - 2\mu_j (\tau \sum_{i:Z_i=j} Y_i + \phi_0 \mu_0) \mu_j \right] \right\} \\
& \Rightarrow \mu_j \mid - \sim N((\tau N_j + \phi_0)^{-1} (\tau \sum_{i:Z_i=j} Y_i + \phi_0 \mu_0), (\tau N_j + \phi_0)^{-1})
\end{aligned}
$$

Task 3

Full conditional of μ_j:

$$
\begin{aligned}
& p(\mu_j \mid Y_{1:n}, Z_{1:n}, \mu_{(-j)}, \mu_0, \phi_0, w_{1:3}, \tau) \\
\propto & \prod_{i=1}^{n} \left(N(Y_i; \mu_{Z_i}, \tau^{-1}) w_{Z_i}\right) p(\mu_{1:3} \mid \mu_0, \phi_0) p(\mu_0) p(\phi_0) p(\tau) \\
\propto & \prod_{i=1}^{n} \left(N(Y_i; \mu_{Z_i}, \tau^{-1})\right) p(\mu_j \mid \mu_0, \phi_0) \\
\propto & \exp\left\{ \frac{-\tau}{2} \sum_{i:Z_i=j} (Y_i - \mu_j)^2 \right\} \exp\{ \frac{-\phi_0}{2} (\mu_j - \mu_0)^2 \} \\
\propto & \exp\left\{ \frac{-1}{2} \left[\mu_j^2 (\tau N_j + \phi_0) - 2\mu_j (\tau \sum_{i:Z_i=j} Y_i + \phi_0 \mu_0) \mu_j \right] \right\} \\
\Rightarrow & \mu_j \mid - \sim N((\tau N_j + \phi_0)^{-1} (\tau \sum_{i:Z_i=j} Y_i + \phi_0 \mu_0), (\tau N_j + \phi_0)^{-1})
\end{aligned}
$$

Task 3

Full conditional for τ:

$$p(\tau \mid Y_{1:n}, Z_{1:n}, \mu_{1:3}, \mu_0, \phi_0, w_{1:3})$$

$$\propto \prod_{i=1}^{n} \left(N(Y_i; \mu_{Z_i}, \tau^{-1}) w_{Z_i} \right) p(\mu_{1:3} \mid \mu_0, \phi_0) p(\mu_0) p(\phi_0) p(\tau)$$

$$\propto \prod_{i=1}^{n} \left(N(Y_i; \mu_{Z_i}, \tau^{-1}) \right) p(\tau)$$

$$\propto \tau^{n/2} \exp\left\{ \frac{-\tau}{2} \sum_{i=1}^{n} (Y_i - \mu_{Z_i})^2 \right\} \tau^{2-1} \exp\{-2\tau\}$$

$$\propto \tau^{2+n/2-1} \exp\{-\tau(2 + \frac{1}{2} \sum_{i=1}^{n} (Y_i - \mu_{Z_i})^2)\}$$

$$\Rightarrow \tau \mid - \sim \text{Gamma}(2 + n/2, 2 + \sum_{i=1}^{n} (Y_i - \mu_{Z_i})^2 / 2)$$

Task 3

Full conditional for τ:

$$
\begin{aligned}
& p(\tau \mid Y_{1:n}, Z_{1:n}, \mu_{1:3}, \mu_0, \phi_0, w_{1:3}) \\
\propto & \prod_{i=1}^{n} \left(N(Y_i; \mu_{Z_i}, \tau^{-1}) w_{Z_i} \right) p(\mu_{1:3} \mid \mu_0, \phi_0) p(\mu_0) p(\phi_0) p(\tau) \\
\propto & \prod_{i=1}^{n} \left(N(Y_i; \mu_{Z_i}, \tau^{-1}) \right) p(\tau) \\
\propto & \tau^{n/2} \exp\left\{ \frac{-\tau}{2} \sum_{i=1}^{n} (Y_i - \mu_{Z_i})^2 \right\} \tau^{2-1} \exp\{-2\tau\} \\
\propto & \tau^{2+n/2-1} \exp\{-\tau(2 + \frac{1}{2} \sum_{i=1}^{n} (Y_i - \mu_{Z_i})^2)\} \\
\Rightarrow & \tau \mid - \sim \text{Gamma}(2 + n/2, 2 + \sum_{i=1}^{n} (Y_i - \mu_{Z_i})^2 / 2)
\end{aligned}
$$

Task 3

Full conditional for τ:

$$
\begin{aligned}
& p(\tau \mid Y_{1:n}, Z_{1:n}, \mu_{1:3}, \mu_0, \phi_0, w_{1:3}) \\
\propto & \prod_{i=1}^{n} \left(N(Y_i; \mu_{Z_i}, \tau^{-1}) w_{Z_i} \right) p(\mu_{1:3} \mid \mu_0, \phi_0) p(\mu_0) p(\phi_0) p(\tau) \\
\propto & \prod_{i=1}^{n} \left(N(Y_i; \mu_{Z_i}, \tau^{-1}) \right) p(\tau) \\
\propto & \tau^{n/2} \exp\left\{ \frac{-\tau}{2} \sum_{i=1}^{n} (Y_i - \mu_{Z_i})^2 \right\} \tau^{2-1} \exp\{-2\tau\} \\
\propto & \tau^{2+n/2-1} \exp\{-\tau(2 + \frac{1}{2} \sum_{i=1}^{n} (Y_i - \mu_{Z_i})^2)\} \\
\Rightarrow & \tau \mid - \sim \text{Gamma}(2 + n/2, 2 + \sum_{i=1}^{n} (Y_i - \mu_{Z_i})^2 / 2)
\end{aligned}
$$

Task 3

Full conditional for τ:

$$
\begin{aligned}
& p(\tau \mid Y_{1:n}, Z_{1:n}, \mu_{1:3}, \mu_0, \phi_0, w_{1:3}) \\
\propto & \prod_{i=1}^{n} \left(N(Y_i; \mu_{Z_i}, \tau^{-1}) w_{Z_i} \right) p(\mu_{1:3} \mid \mu_0, \phi_0) p(\mu_0) p(\phi_0) p(\tau) \\
\propto & \prod_{i=1}^{n} \left(N(Y_i; \mu_{Z_i}, \tau^{-1}) \right) p(\tau) \\
\propto & \tau^{n/2} \exp\left\{ \frac{-\tau}{2} \sum_{i=1}^{n} (Y_i - \mu_{Z_i})^2 \right\} \tau^{2-1} \exp\{-2\tau\} \\
\propto & \tau^{2+n/2-1} \exp\{-\tau(2 + \frac{1}{2} \sum_{i=1}^{n} (Y_i - \mu_{Z_i})^2)\} \\
\Rightarrow & \tau \mid - \sim \text{Gamma}(2 + n/2, 2 + \sum_{i=1}^{n} (Y_i - \mu_{Z_i})^2 / 2)
\end{aligned}
$$

Task 3

Full conditional for τ:

$$
\begin{aligned}
&p(\tau \mid Y_{1:n}, Z_{1:n}, \mu_{1:3}, \mu_0, \phi_0, w_{1:3}) \\
\propto &\prod_{i=1}^{n} \left(N(Y_i; \mu_{Z_i}, \tau^{-1}) w_{Z_i} \right) p(\mu_{1:3} \mid \mu_0, \phi_0) p(\mu_0) p(\phi_0) p(\tau) \\
\propto &\prod_{i=1}^{n} \left(N(Y_i; \mu_{Z_i}, \tau^{-1}) \right) p(\tau) \\
\propto &\tau^{n/2} \exp\left\{ \frac{-\tau}{2} \sum_{i=1}^{n} (Y_i - \mu_{Z_i})^2 \right\} \tau^{2-1} \exp\{-2\tau\} \\
\propto &\tau^{2+n/2-1} \exp\{-\tau(2 + \frac{1}{2} \sum_{i=1}^{n} (Y_i - \mu_{Z_i})^2)\}
\end{aligned}
$$

$$
\Rightarrow \tau \mid - \sim \text{Gamma}(2 + n/2, 2 + \sum_{i=1}^{n} (Y_i - \mu_{Z_i})^2/2)
$$

Task 3

Full conditional for τ:

$$
\begin{aligned}
& p(\tau \mid Y_{1:n}, Z_{1:n}, \mu_{1:3}, \mu_0, \phi_0, w_{1:3}) \\
\propto & \prod_{i=1}^{n} \left(N(Y_i; \mu_{Z_i}, \tau^{-1}) w_{Z_i} \right) p(\mu_{1:3} \mid \mu_0, \phi_0) p(\mu_0) p(\phi_0) p(\tau) \\
\propto & \prod_{i=1}^{n} \left(N(Y_i; \mu_{Z_i}, \tau^{-1}) \right) p(\tau) \\
\propto & \tau^{n/2} \exp\left\{ \frac{-\tau}{2} \sum_{i=1}^{n} (Y_i - \mu_{Z_i})^2 \right\} \tau^{2-1} \exp\{-2\tau\} \\
\propto & \tau^{2+n/2-1} \exp\{-\tau(2 + \frac{1}{2} \sum_{i=1}^{n} (Y_i - \mu_{Z_i})^2)\} \\
\Rightarrow & \tau \mid - \sim \text{Gamma}(2 + n/2, 2 + \sum_{i=1}^{n} (Y_i - \mu_{Z_i})^2 / 2)
\end{aligned}
$$

Task 3

Full conditional for Z_i:

$$
\begin{aligned}
& p(Z_i \mid Y_{1:n}, Z_{1:n}, \mu_{1:3}, \mu_0, \phi_0, w_{1:3}, \tau) \\
\propto & \prod_{i=1}^{n} \left(N(Y_i; \mu_{Z_i}, \tau^{-1}) w_{Z_i} \right) p(\mu_{1:3} \mid \mu_0, \phi_0) p(\mu_0) p(\phi_0) p(\tau) \\
\propto & N(Y_i; \mu_{Z_i}, \tau^{-1}) w_{Z_i} \\
\Rightarrow & P(Z_i = j \mid -) \propto w_j N(Y_i; \mu_j, \tau^{-1}).
\end{aligned}
$$

Task 3

Full conditional for Z_i:

$$
\begin{aligned}
& p(Z_i \mid Y_{1:n}, Z_{1:n}, \mu_{1:3}, \mu_0, \phi_0, w_{1:3}, \tau) \\
\propto & \prod_{i=1}^{n} \left(N(Y_i; \mu_{Z_i}, \tau^{-1}) w_{Z_i} \right) p(\mu_{1:3} \mid \mu_0, \phi_0) p(\mu_0) p(\phi_0) p(\tau) \\
\propto & N(Y_i; \mu_{Z_i}, \tau^{-1}) w_{Z_i} \\
\Rightarrow & P(Z_i = j \mid -) \propto w_j N(Y_i; \mu_j, \tau^{-1}).
\end{aligned}
$$

Task 3

Full conditional for Z_i:

$$
\begin{aligned}
&p(Z_i \mid Y_{1:n}, Z_{1:n}, \mu_{1:3}, \mu_0, \phi_0, w_{1:3}, \tau) \\
\propto &\prod_{i=1}^{n} \left(N(Y_i; \mu_{Z_i}, \tau^{-1}) w_{Z_i} \right) p(\mu_{1:3} \mid \mu_0, \phi_0) p(\mu_0) p(\phi_0) p(\tau) \\
\propto &N(Y_i; \mu_{Z_i}, \tau^{-1}) w_{Z_i} \\
\Rightarrow &P(Z_i = j \mid -) \propto w_j N(Y_i; \mu_j, \tau^{-1}).
\end{aligned}
$$

Task 3

Full conditional for Z_i:

$$
\begin{aligned}
&p(Z_i \mid Y_{1:n}, Z_{1:n}, \mu_{1:3}, \mu_0, \phi_0, w_{1:3}, \tau) \\
\propto &\prod_{i=1}^{n} \left(N(Y_i; \mu_{Z_i}, \tau^{-1}) w_{Z_i} \right) p(\mu_{1:3} \mid \mu_0, \phi_0) p(\mu_0) p(\phi_0) p(\tau) \\
\propto &N(Y_i; \mu_{Z_i}, \tau^{-1}) w_{Z_i} \\
\Rightarrow &P(Z_i = j \mid -) \propto w_j N(Y_i; \mu_j, \tau^{-1}).
\end{aligned}
$$

Task 3

Summary:

$$p(\mu_0 \mid \ldots) = N\left(\frac{\sigma_0^2 \sum_{i=1}^3 \mu_i}{1/3 + 3\sigma_0^{-2}}, (1/3 + 3\sigma_0^{-2})^{-1}\right),$$

$$p(\sigma_0^2 \mid \ldots) = IG\left(2 + 3/2, 2 + (1/2)\sum_{i=1}^{3}(\mu_i - \mu_0)^2\right),$$

$$p(\epsilon^2 \mid \ldots) = IG\left(2 + n/2, 2 + (1/2)\sum_{i=1}^{n}(Y_i - \mu_{Z_i})^2\right),$$

$$p(\boldsymbol{w} \mid \ldots) = Dir(3, (1 + N_1, 1 + N_2, 1 + N_3)),$$

$$p(\mu_j \mid \ldots) = N\left(\left(\mu_0\sigma_0^{-2} + \epsilon^{-2}\sum_{i:Z_i=j} y_i\right)(\sigma_0^{-2} + N_j\epsilon^{-2})^{-1}, \left(\sigma_0^{-2} + N_j\epsilon^{-2}\right)^{-1}\right),$$

$$P(Z_i = j) = \frac{wjN(y_i \mid \mu_j, \epsilon^2)}{\sum_{k=1}^3 w_k N(y_i \mid \mu_k, \epsilon^2)}.$$

Load Libraries

```r
# load libraries #
library(gtools)
library(pscl)
```

```
## Classes and Methods for R developed in the
## Political Science Computational Laboratory
## Department of Political Science
## Stanford University
## Simon Jackman
## hurdle and zeroinfl functions by Achim Zeileis
```

Load Data

```r
# set seed for reproducability #
set.seed(666)
# set number of simulations #
nsims = 10
# read in the data #
y = read.csv("Lab8Mixture.csv", header=FALSE)
```

Prepare Data

```r
# set matrix for w values #
w = matrix(0, nrow = nsims, ncol=3)
colnames(w) = c("w1", "w2", "w3")
```

Initialization

```r
# set inital values for w #
w[1,] = c(1/3,1/3,1/3)
# set matrix for w values #
mu_j = matrix(0, nrow = nsims, ncol=3)
colnames(mu_j) = c("mu1", "mu2", "mu3")
# set inital values for w #
mu_j[1,] = c(1,1,1)
```

Initialization

```
# create vectors #
sigma0 = c()
e2 = c()
mu0 = c()
# set inital values #
sigma0[1] = 1
e2[1] = 1
mu0[1] = 1
```

Initialization

```r
# create matrix for z values #
z = matrix(0, ncol = nrow(y), nrow = nsims)
# set inital values for z #
z[1,] = sample(c(1,2,3), size = nrow(y), replace = TRUE)
```

Initialize Gibbs

```r
N_vals = c()
S_vals = c()
Y_vals = list()
```

Gibbs Sampler (not elegant)

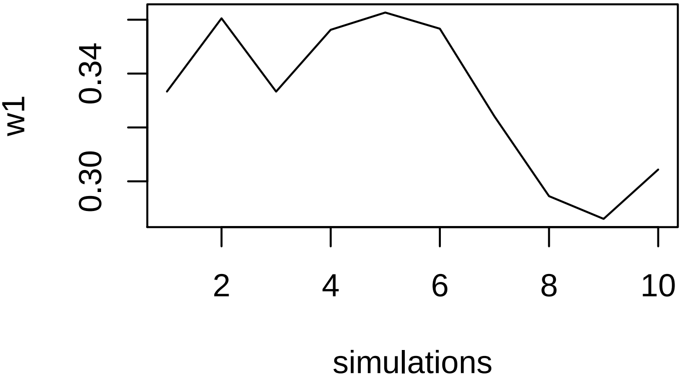
```
# gibbs sampling #
for (sim in 2:nsims) {

  for (i in 1:3) { # find values for each category
    N_vals[i] = sum(z[sim-1, ] == i)
    Y_vals[[i]] = which(z[sim-1, ] == i)
    S_vals[i] = sum((y[Y_vals[[i]], ] - mu_j[sim-1, i])^2)
  }

  w[sim,] = rdirichlet(n = 1, alpha = N_vals+1)  # sample w values

  e2[sim] = rigamma(n = 1, alpha = (4+sum(N_vals))/2, # sample e2 values
                    beta = sum(S_vals)/2)
  # sample mu0 values
  mu0[sim] = rnorm(n = 1,
              mean = ((1/((1/3) + (3 / sigma0[sim-1])))*
                     (sum(mu_j[sim-1, ]) / sigma0[sim-1])),
              sd = sqrt((1/((1/3) + (3 / sigma0[sim-1])))))
  sigma0[sim] = rigamma(n = 1, alpha = 3.5, # sample sigma0 values
                        beta = (2 + ((sum((mu_j[sim-1, ] - mu0[sim])^2))/2)))
  for (g in 1:3) { # sample mu values
    mu_j[sim, g] = rnorm(n = 1, mean = (((mu0[sim]/sigma0[sim]) +
                                          (sum(y[Y_vals[[g]], ])/e2[sim])) *
                                         (1 / ((1/sigma0[sim]) + (N_vals[g]/e2[sim])))),
                 sd = sqrt((1 / ((1/sigma0[sim]) + (N_vals[g]/e2[sim])))))
  }
  # sample z values
  for(h in 1:nrow(y)) {
    prob_vals = w[sim, ]*dnorm(y[h,], mean = mu_j[sim, ], sd = sqrt(e2[sim]))
    z[sim, h] = sample(x = c(1,2,3), size = 1, replace = TRUE,
                       prob =  prob_vals/sum(prob_vals))
  }
}
mixture_results = cbind(w, mu_j, e2, mu0, sigma0, z)
```

Task 5

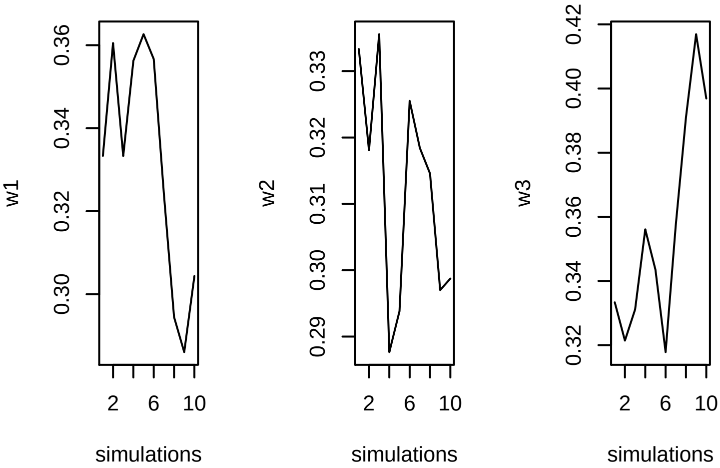
Now, we want to create trace plots, averages, and 95% confidence intervals for the marginal posterior distributions of all the parameters.

Traceplot of w_1

```
plot(1:nrow(mixture_results), mixture_results[,1],
      type = "l", xlab = "simulations",
      ylab = colnames(mixture_results)[1])
```



Traceplots of w_1, w_2, w_3



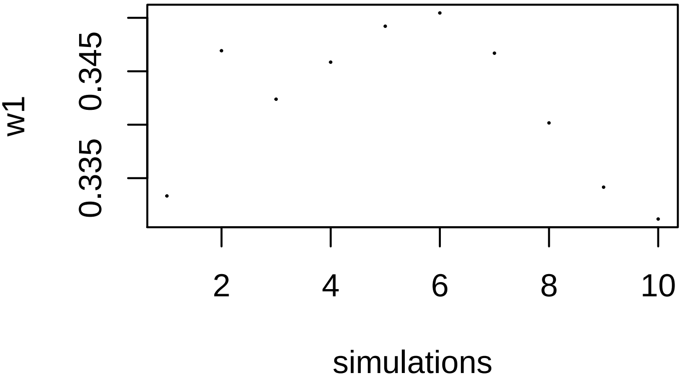
Traceplots of weights

What can you conclude regarding the traceplot of the weights?

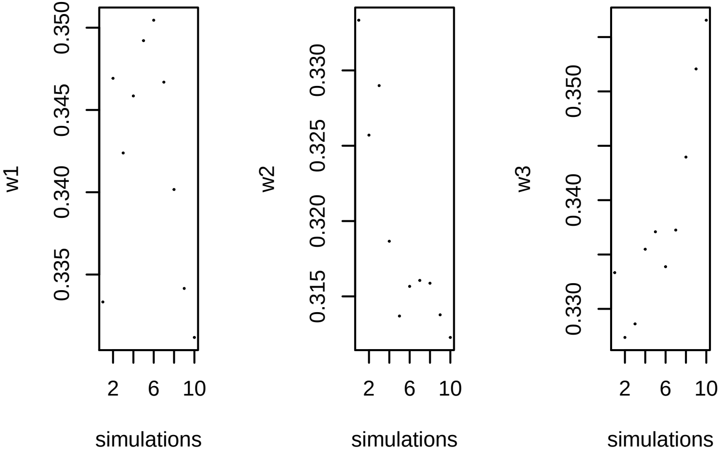
What should you do to fix the situation?

Running average plots for w_1

```
# plot moving average trace plots
  mov_avg = cumsum(mixture_results[,1])/(1:nrow(mixture_results))
  plot(1:length(mov_avg), mov_avg, cex = 0.1,
       xlab = "simulations", ylab = colnames(mixture_results)[1])
```



Running average plots of w_1, w_2, w_3



Calculate our 95% credible intervals

```r
# calculate 95% credible intervals #
for (i in 1:9) {
  upper = quantile(mixture_results[,i], 0.975)
  lower = quantile(mixture_results[,i], 0.25)
  cat("The 95% credible interval for",
      colnames(mixture_results)[i], "is between",
      lower, "and", upper, "\n")
}
```

```
## The 95% credible interval for w1 is between 0.3093064 and 0.3621755
## The 95% credible interval for w2 is between 0.2974454 and 0.3350666
## The 95% credible interval for w3 is between 0.3316713 and 0.4124205
## The 95% credible interval for mu1 is between 2.448104 and 3.403709
## The 95% credible interval for mu2 is between 3.008058 and 3.573371
## The 95% credible interval for mu3 is between 2.156434 and 2.767639
## The 95% credible interval for e2 is between 4.059747 and 6.904096
## The 95% credible interval for mu0 is between 1.875007 and 3.454596
## The 95% credible interval for sigma0 is between 0.5745476 and 2.452617
```