

## 1: The Bisection Method (25%)

For function  $f(x) = x^3 + 6x^2 + \pi x - 12$ , the derivative is  $f'(x) = 3x^2 + 12x + \pi$ . The we can calculate that zeros of the derivative are  $\frac{-12-\sqrt{12(12-\pi)}}{6}$  and  $\frac{-12+\sqrt{12(12-\pi)}}{6}$ .

$f(\frac{-12-\sqrt{12(12-\pi)}}{6}) = 7.864841$  and  $f(\frac{-12+\sqrt{12(12-\pi)}}{6}) = -12.43121$  Hence, the function  $f$  has totally 3 zeros.

**Algorithm:** Bisection Method in the R file.

**Result:** zeros -4.837944, -2.259727, and 1.097664.

## 2: Poisson Regression - Newton's Method (25%)

(1) Since  $y_i \sim \text{Poisson}(\lambda_i)$  and  $\log(\lambda_i) = \alpha + \beta x_i + \gamma x_i^2$ , we can get the Likelihood function:

$$L(\alpha, \beta, \gamma | \mathbf{x}, \mathbf{y}) = \prod_{i=1}^n \frac{\lambda_i^{y_i} e^{-\lambda_i}}{y_i!} = \prod_{i=1}^n \frac{e^{(\alpha + \beta x_i + \gamma x_i^2)y_i} e^{-e^{\alpha + \beta x_i + \gamma x_i^2}}}{y_i!}$$

(2) The log-Likelihood function is

$$l(\alpha, \beta, \gamma | \mathbf{x}, \mathbf{y}) = \sum_{i=1}^n (\alpha + \beta x_i + \gamma x_i^2) y_i - e^{\alpha + \beta x_i + \gamma x_i^2} - \log y_i!$$

Then we have:

$$\begin{aligned} \left. \frac{\partial l(\alpha, \beta, \gamma | \mathbf{x}, \mathbf{y})}{\partial \alpha} \right|_{\hat{\alpha}} &= \sum_{i=1}^n [y_i - e^{\alpha + \beta x_i + \gamma x_i^2}] \Big|_{\hat{\alpha}} = 0 \\ \left. \frac{\partial l(\alpha, \beta, \gamma | \mathbf{x}, \mathbf{y})}{\partial \beta} \right|_{\hat{\beta}} &= \sum_{i=1}^n [x_i y_i - x_i e^{\alpha + \beta x_i + \gamma x_i^2}] \Big|_{\hat{\beta}} = 0 \\ \left. \frac{\partial l(\alpha, \beta, \gamma | \mathbf{x}, \mathbf{y})}{\partial \gamma} \right|_{\hat{\gamma}} &= \sum_{i=1}^n [x_i^2 y_i - x_i^2 e^{\alpha + \beta x_i + \gamma x_i^2}] \Big|_{\hat{\gamma}} = 0 \end{aligned}$$

$$\text{Let } \mathbf{F}(\mathbf{x}) = \begin{pmatrix} \frac{\partial l(\alpha, \beta, \gamma | \mathbf{x}, \mathbf{y})}{\partial \alpha} \\ \frac{\partial l(\alpha, \beta, \gamma | \mathbf{x}, \mathbf{y})}{\partial \beta} \\ \frac{\partial l(\alpha, \beta, \gamma | \mathbf{x}, \mathbf{y})}{\partial \gamma} \end{pmatrix}, \text{ then } \mathbf{F}'(\mathbf{x}) = \begin{pmatrix} \frac{\partial^2 l(\alpha, \beta, \gamma | \mathbf{x}, \mathbf{y})}{\partial \alpha^2} & \frac{\partial^2 l(\alpha, \beta, \gamma | \mathbf{x}, \mathbf{y})}{\partial \beta \partial \alpha} & \frac{\partial^2 l(\alpha, \beta, \gamma | \mathbf{x}, \mathbf{y})}{\partial \gamma \partial \alpha} \\ \frac{\partial^2 l(\alpha, \beta, \gamma | \mathbf{x}, \mathbf{y})}{\partial \alpha \partial \beta} & \frac{\partial^2 l(\alpha, \beta, \gamma | \mathbf{x}, \mathbf{y})}{\partial \beta^2} & \frac{\partial^2 l(\alpha, \beta, \gamma | \mathbf{x}, \mathbf{y})}{\partial \gamma \partial \beta} \\ \frac{\partial^2 l(\alpha, \beta, \gamma | \mathbf{x}, \mathbf{y})}{\partial \alpha \partial \gamma} & \frac{\partial^2 l(\alpha, \beta, \gamma | \mathbf{x}, \mathbf{y})}{\partial \beta \partial \gamma} & \frac{\partial^2 l(\alpha, \beta, \gamma | \mathbf{x}, \mathbf{y})}{\partial \gamma^2} \end{pmatrix}, \text{ in}$$

which

$$\begin{aligned}
\frac{\partial^2 l(\alpha, \beta, \gamma | \mathbf{x}, \mathbf{y})}{\partial \alpha^2} &= \sum_{i=1}^n -e^{\alpha + \beta x_i + \gamma x_i^2} \\
\frac{\partial^2 l(\alpha, \beta, \gamma | \mathbf{x}, \mathbf{y})}{\partial \alpha \partial \beta} &= \frac{\partial^2 l(\alpha, \beta, \gamma | \mathbf{x}, \mathbf{y})}{\partial \beta \partial \alpha} = \sum_{i=1}^n -x_i e^{\alpha + \beta x_i + \gamma x_i^2} \\
\frac{\partial^2 l(\alpha, \beta, \gamma | \mathbf{x}, \mathbf{y})}{\partial \alpha \partial \gamma} &= \frac{\partial^2 l(\alpha, \beta, \gamma | \mathbf{x}, \mathbf{y})}{\partial \gamma \partial \alpha} = \sum_{i=1}^n -x_i^2 e^{\alpha + \beta x_i + \gamma x_i^2} \\
\frac{\partial^2 l(\alpha, \beta, \gamma | \mathbf{x}, \mathbf{y})}{\partial \beta^2} &= \sum_{i=1}^n -x_i^2 e^{\alpha + \beta x_i + \gamma x_i^2} \\
\frac{\partial^2 l(\alpha, \beta, \gamma | \mathbf{x}, \mathbf{y})}{\partial \beta \partial \gamma} &= \frac{\partial^2 l(\alpha, \beta, \gamma | \mathbf{x}, \mathbf{y})}{\partial \gamma \partial \beta} = \sum_{i=1}^n -x_i^3 e^{\alpha + \beta x_i + \gamma x_i^2} \\
\frac{\partial^2 l(\alpha, \beta, \gamma | \mathbf{x}, \mathbf{y})}{\partial \gamma^2} &= \sum_{i=1}^n -x_i^4 e^{\alpha + \beta x_i + \gamma x_i^2}
\end{aligned}$$

Therefore, by Newton's Method, given initial guess  $\alpha^{(0)}$ ,  $\beta^{(0)}$ , and  $\gamma^{(0)}$ , for each iteration:

$$\begin{pmatrix} \alpha^{(n)} \\ \beta^{(n)} \\ \gamma^{(n)} \end{pmatrix} = \begin{pmatrix} \alpha^{(n-1)} \\ \beta^{(n-1)} \\ \gamma^{(n-1)} \end{pmatrix} - \mathbf{F}'[(\mathbf{x})]^{-1} \mathbf{F}(\mathbf{x}).$$

Hence, for the algorithm:  $\mathbf{x}^{(n)} = \begin{pmatrix} \alpha^{(n)} \\ \beta^{(n)} \\ \gamma^{(n)} \end{pmatrix}$ ,

STEP 1: Solve  $\mathbf{F}'(\mathbf{x}^{(n)}) \Delta \mathbf{x}^{(n)} = -\mathbf{F}(\mathbf{x}^{(n)})$ ;

STEP 2: Update by  $\mathbf{x}^{(n+1)} = \mathbf{x}^{(n)} + \Delta \mathbf{x}^{(n)}$ .

(3)

**Algorithm:** Newton's Method code in the R file.

**Result:**  $\alpha = 1.503533$ ,  $\beta = 1.052351$ , and  $\gamma = 1.957396$ .

---

### 3: Logistic Regression - Newton's Method (20%)

---

(1) Since  $y_i \sim \text{Bernoulli}(p_i)$  and  $\text{logit}(p_i) = \alpha + \beta x_i$ , we can know that  $f(y_i, p_i) = p_i^{y_i} (1 - p_i)^{1-y_i}$ , and  $p_i = \frac{e^{\alpha + \beta x_i}}{1 + e^{\alpha + \beta x_i}}$ . Then we can get the following Likelihood function:

$$\begin{aligned}
L(\alpha, \beta | \mathbf{x}, \mathbf{y}) &= \prod_{i=1}^n p_i^{y_i} (1 - p_i)^{1-y_i} \\
&= \prod_{i=1}^n \left( \frac{e^{\alpha + \beta x_i}}{1 + e^{\alpha + \beta x_i}} \right)^{y_i} \left( 1 - \frac{e^{\alpha + \beta x_i}}{1 + e^{\alpha + \beta x_i}} \right)^{1-y_i}
\end{aligned}$$

(2) The log-Likelihood function is

$$\begin{aligned} l(\alpha, \beta | \mathbf{x}, \mathbf{y}) &= \sum_{i=1}^n y_i (\alpha + \beta x_i - \log(1 + e^{\alpha + \beta x_i})) + (1 - y_i) \log\left(\frac{1}{1 + e^{\alpha + \beta x_i}}\right) \\ &= \sum_{i=1}^n \alpha x_i + \beta x_i y_i - \log(1 + e^{\alpha + \beta x_i}) \end{aligned}$$

Then we have:

$$\begin{aligned} \left. \frac{\partial l(\alpha, \beta | \mathbf{x}, \mathbf{y})}{\partial \alpha} \right|_{\hat{\alpha}} &= \sum_{i=1}^n \left[ y_i - \frac{e^{\alpha + \beta x_i}}{1 + e^{\alpha + \beta x_i}} \right] \Big|_{\hat{\alpha}} = 0 \\ \left. \frac{\partial l(\alpha, \beta | \mathbf{x}, \mathbf{y})}{\partial \beta} \right|_{\hat{\beta}} &= \sum_{i=1}^n \left[ x_i y_i - \frac{x_i e^{\alpha + \beta x_i}}{1 + e^{\alpha + \beta x_i}} \right] \Big|_{\hat{\beta}} = 0 \end{aligned}$$

Let  $\mathbf{F}(\mathbf{x}) = \begin{pmatrix} \frac{\partial l(\alpha, \beta | \mathbf{x}, \mathbf{y})}{\partial \alpha} \\ \frac{\partial l(\alpha, \beta | \mathbf{x}, \mathbf{y})}{\partial \beta} \end{pmatrix}$ , then  $\mathbf{F}'(\mathbf{x}) = \begin{pmatrix} \frac{\partial^2 l(\alpha, \beta | \mathbf{x}, \mathbf{y})}{\partial \alpha^2} & \frac{\partial^2 l(\alpha, \beta | \mathbf{x}, \mathbf{y})}{\partial \beta \partial \alpha} \\ \frac{\partial^2 l(\alpha, \beta | \mathbf{x}, \mathbf{y})}{\partial \alpha \partial \beta} & \frac{\partial^2 l(\alpha, \beta | \mathbf{x}, \mathbf{y})}{\partial \beta^2} \end{pmatrix}$ , in which

$$\begin{aligned} \frac{\partial^2 l(\alpha, \beta | \mathbf{x}, \mathbf{y})}{\partial \alpha^2} &= -\frac{e^{\alpha + \beta x_i}}{(1 + e^{\alpha + \beta x_i})^2} \\ \frac{\partial^2 l(\alpha, \beta | \mathbf{x}, \mathbf{y})}{\partial \beta \partial \alpha} &= \frac{\partial^2 l(\alpha, \beta | \mathbf{x}, \mathbf{y})}{\partial \alpha \partial \beta} = -\frac{x_i e^{\alpha + \beta x_i}}{(1 + e^{\alpha + \beta x_i})^2} \\ \frac{\partial^2 l(\alpha, \beta | \mathbf{x}, \mathbf{y})}{\partial \beta^2} &= -\frac{x_i^2 e^{\alpha + \beta x_i}}{(1 + e^{\alpha + \beta x_i})^2} \end{aligned}$$

Therefore, by Newton's Method, given initial guess  $\alpha^{(0)}$  and  $\beta^{(0)}$ , for each iteration:

$$\begin{pmatrix} \alpha^{(n)} \\ \beta^{(n)} \end{pmatrix} = \begin{pmatrix} \alpha^{(n-1)} \\ \beta^{(n-1)} \end{pmatrix} - \mathbf{F}'(\mathbf{x})^{-1} \mathbf{F}(\mathbf{x}).$$

Hence, for the algorithm:  $\mathbf{x}^{(n)} = \begin{pmatrix} \alpha^{(n)} \\ \beta^{(n)} \end{pmatrix}$ ,

STEP 1: Solve  $\mathbf{F}'(\mathbf{x}^{(n)}) \Delta \mathbf{x}^{(n)} = -\mathbf{F}(\mathbf{x}^{(n)})$ ;

STEP 2: Update by  $\mathbf{x}^{(n+1)} = \mathbf{x}^{(n)} + \Delta \mathbf{x}^{(n)}$ .

(3)

**Algorithm:** Newton's Method code in the R file.

**Result:**  $\alpha = 1.564284$  and  $\beta = 1.771093$ .

---

#### 4: EM Algorithm (30%)

---

(1) Observed data:  $Y_i$  for  $i = 1, 2, \dots, 8000$ ; Missing data:  $Z_i$  for  $i = 1, 2, \dots, 8000$ , where  $Z_i = 1, 2$ , or  $3$  for low, middle, and high income respectively.

Since  $Y_i|(Z_i = k) \sim N(\mu_k, \sigma_k^2)$ , with proportion  $\pi_k$  ( $\pi_3 = 1 - (\pi_1 + \pi_2)$ ), we can formulate the **complete – data Likelihood function** as:

$$L(\pi_1, \pi_2, \mu_1, \mu_2, \mu_3, \sigma_1, \sigma_2, \sigma_3 | \mathbf{Y}, \mathbf{Z}) \\ = \prod_{i=1}^n \left[ \pi_1 \frac{1}{\sqrt{2\pi}\sigma_1} e^{-\frac{(y_i - \mu_1)^2}{2\sigma_1^2}} \right]^{I(Z_i=1)} \left[ \pi_2 \frac{1}{\sqrt{2\pi}\sigma_2} e^{-\frac{(y_i - \mu_2)^2}{2\sigma_2^2}} \right]^{I(Z_i=2)} \left[ (1 - \pi_1 - \pi_2) \frac{1}{\sqrt{2\pi}\sigma_3} e^{-\frac{(y_i - \mu_3)^2}{2\sigma_3^2}} \right]^{I(Z_i=3)}$$

The observed-data Likelihood function is:

$$L(\pi_1, \pi_2, \mu_1, \mu_2, \mu_3, \sigma_1, \sigma_2, \sigma_3 | \mathbf{Y}) \\ = \prod_{i=1}^n \left[ \left[ \pi_1 \frac{1}{\sqrt{2\pi}\sigma_1} e^{-\frac{(y_i - \mu_1)^2}{2\sigma_1^2}} \right] + \left[ \pi_2 \frac{1}{\sqrt{2\pi}\sigma_2} e^{-\frac{(y_i - \mu_2)^2}{2\sigma_2^2}} \right] + \left[ (1 - \pi_1 - \pi_2) \frac{1}{\sqrt{2\pi}\sigma_3} e^{-\frac{(y_i - \mu_3)^2}{2\sigma_3^2}} \right] \right]$$

(2) The complete-data log-Likelihood function is:

$$l(\pi_1, \pi_2, \mu_1, \mu_2, \mu_3, \sigma_1, \sigma_2, \sigma_3 | \mathbf{Y}, \mathbf{Z}) = \sum_{i=1}^n \left[ I(Z_i = 1) \left[ \log \pi_1 - \log \sqrt{2\pi} - \frac{1}{2} \log \sigma_1^2 - \frac{(y_i - \mu_1)^2}{2\sigma_1^2} \right] \right. \\ \left. + I(Z_i = 2) \left[ \log \pi_2 - \log \sqrt{2\pi} - \frac{1}{2} \log \sigma_2^2 - \frac{(y_i - \mu_2)^2}{2\sigma_2^2} \right] \right. \\ \left. + I(Z_i = 3) \left[ \log(1 - \pi_1 - \pi_2) - \log \sqrt{2\pi} - \frac{1}{2} \log \sigma_3^2 - \frac{(y_i - \mu_3)^2}{2\sigma_3^2} \right] \right]$$

Then we can calculate:

$$\begin{aligned} \left. \frac{\partial l}{\partial \pi_1} \right|_{\hat{\pi}_1, \hat{\pi}_2} &= \sum_{i=1}^n \left[ I(Z_i = 1) \frac{1}{\pi_1} - I(Z_i = 3) \frac{1}{1 - \pi_1 - \pi_2} \right] \Big|_{\hat{\pi}_1, \hat{\pi}_2} = 0 \\ \left. \frac{\partial l}{\partial \pi_2} \right|_{\hat{\pi}_1, \hat{\pi}_2} &= \sum_{i=1}^n \left[ I(Z_i = 2) \frac{1}{\pi_2} - I(Z_i = 3) \frac{1}{1 - \pi_1 - \pi_2} \right] \Big|_{\hat{\pi}_1, \hat{\pi}_2} = 0 \\ \left. \frac{\partial l}{\partial \mu_1} \right|_{\hat{\mu}_1, \hat{\sigma}_1^2} &= \sum_{i=1}^n \left[ I(Z_i = 1) \frac{y_i - \mu_1}{\sigma_1^2} \right] \Big|_{\hat{\mu}_1, \hat{\sigma}_1^2} = 0 \\ \left. \frac{\partial l}{\partial \mu_2} \right|_{\hat{\mu}_2, \hat{\sigma}_2^2} &= \sum_{i=1}^n \left[ I(Z_i = 2) \frac{y_i - \mu_2}{\sigma_2^2} \right] \Big|_{\hat{\mu}_2, \hat{\sigma}_2^2} = 0 \\ \left. \frac{\partial l}{\partial \mu_3} \right|_{\hat{\mu}_3, \hat{\sigma}_3^2} &= \sum_{i=1}^n \left[ I(Z_i = 3) \frac{y_i - \mu_3}{\sigma_3^2} \right] \Big|_{\hat{\mu}_3, \hat{\sigma}_3^2} = 0 \\ \left. \frac{\partial l}{\partial \sigma_1^2} \right|_{\hat{\mu}_1, \hat{\sigma}_1^2} &= \sum_{i=1}^n \left[ I(Z_i = 1) \left( -\frac{1}{2} \frac{1}{\sigma_1^2} + \frac{(y_i - \mu_1)^2}{2(\sigma_1^2)^2} \right) \right] \Big|_{\hat{\mu}_1, \hat{\sigma}_1^2} = 0 \\ \left. \frac{\partial l}{\partial \sigma_2^2} \right|_{\hat{\mu}_2, \hat{\sigma}_2^2} &= \sum_{i=1}^n \left[ I(Z_i = 2) \left( -\frac{1}{2} \frac{1}{\sigma_2^2} + \frac{(y_i - \mu_2)^2}{2(\sigma_2^2)^2} \right) \right] \Big|_{\hat{\mu}_2, \hat{\sigma}_2^2} = 0 \\ \left. \frac{\partial l}{\partial \sigma_3^2} \right|_{\hat{\mu}_3, \hat{\sigma}_3^2} &= \sum_{i=1}^n \left[ I(Z_i = 3) \left( -\frac{1}{2} \frac{1}{\sigma_3^2} + \frac{(y_i - \mu_3)^2}{2(\sigma_3^2)^2} \right) \right] \Big|_{\hat{\mu}_3, \hat{\sigma}_3^2} = 0 \end{aligned}$$

Hence, we can conclude that:

$$\begin{aligned}
\hat{\pi}_1 &= \frac{\sum_{i=1}^n I(Z_i = 1)(1 - \hat{\pi}_2)}{\sum_{i=1}^n [I(Z_i = 1) + I(Z_i = 3)]} \\
\hat{\pi}_2 &= \frac{\sum_{i=1}^n I(Z_i = 2)(1 - \hat{\pi}_1)}{\sum_{i=1}^n [I(Z_i = 2) + I(Z_i = 3)]} \\
\hat{\mu}_1 &= \frac{\sum_{i=1}^n I(Z_i = 1)y_i}{\sum_{i=1}^n I(Z_i = 1)} \\
\hat{\mu}_2 &= \frac{\sum_{i=1}^n I(Z_i = 2)y_i}{\sum_{i=1}^n I(Z_i = 2)} \\
\hat{\mu}_3 &= \frac{\sum_{i=1}^n I(Z_i = 3)y_i}{\sum_{i=1}^n I(Z_i = 3)} \\
\hat{\sigma}_1^2 &= \frac{\sum_{i=1}^n I(Z_i = 1)(y_i^2 - \hat{\mu}_1^2 + 2\hat{\mu}_1 y_i)}{\sum_{i=1}^n I(Z_i = 1)} \\
\hat{\sigma}_2^2 &= \frac{\sum_{i=1}^n I(Z_i = 2)(y_i^2 - \hat{\mu}_2^2 + 2\hat{\mu}_2 y_i)}{\sum_{i=1}^n I(Z_i = 2)} \\
\hat{\sigma}_3^2 &= \frac{\sum_{i=1}^n I(Z_i = 3)(y_i^2 - \hat{\mu}_3^2 + 2\hat{\mu}_3 y_i)}{\sum_{i=1}^n I(Z_i = 3)}
\end{aligned}$$

Then given initial guess  $\pi_1^{(0)}, \pi_2^{(0)}, \mu_1^{(0)}, \mu_2^{(0)}, \sigma_1^{(0)}, \sigma_2^{(0)}$ , we define a  $Q$  function by  $Q(\Theta; \Theta^{(t)}) = E_{\Theta^{(t)}}(l(\Theta | \mathbf{Y}, \mathbf{Z}) | \mathbf{Y})$ :

$$\begin{aligned}
Q(\Theta; \Theta^{(t)}) &= E_{\pi_1^{(t)}, \pi_2^{(t)}, \mu_1^{(t)}, \mu_2^{(t)}, \mu_3^{(t)}, \sigma_1^{2(t)}, \sigma_2^{2(t)}, \sigma_3^{2(t)}}(l(\pi_1, \pi_2, \mu_1, \mu_2, \mu_3, \sigma_1, \sigma_2, \sigma_3 | \mathbf{Y}, \mathbf{Z}) | \mathbf{Y}) \\
&= \sum_{i=1}^n \left[ I(Z_i = 1) \left[ \log \pi_1^{(t)} - \log \sqrt{2\pi} - \frac{1}{2} \log \sigma_1^{2(t)} - \frac{(y_i - \mu_1^{(t)})^2}{2\sigma_1^{2(t)}} \right] \right. \\
&\quad + I(Z_i = 2) \left[ \log \pi_2^{(t)} - \log \sqrt{2\pi} - \frac{1}{2} \log \sigma_2^{2(t)} - \frac{(y_i - \mu_2^{(t)})^2}{2\sigma_2^{2(t)}} \right] \\
&\quad \left. + I(Z_i = 3) \left[ \log(1 - \pi_1^{(t)} - \pi_2^{(t)}) - \log \sqrt{2\pi} - \frac{1}{2} \log \sigma_3^{2(t)} - \frac{(y_i - \mu_3^{(t)})^2}{2\sigma_3^{2(t)}} \right] \right]
\end{aligned}$$

(3)

(4)