LOGISTIC REGRESSION FOR KYPHOSIS

The data for 83 children who have had spine surgery consists of

Kyphosis	Binary response registering the presence of kyphosis
Age	Age of the patient in months
Number	The number of vertebra involved in surgery
Start	The number of the initial vertebra in surgery

The *logistic regression model* is given by

$$P(Y = 1 | X_1 = x_1, \ldots, X_m = x_m) = \frac{\exp(\beta_0 + \sum_{k=1}^{n} \beta_k x_k)}{1 + \exp(\beta_0 + \sum_{k=1}^{n} \beta_k x_k)},$$

where $\beta_0, \ldots, \beta_m$ are parameters, Y is the 0/1 response and $X_1, \ldots, X_m$ are covariates.

a. Write down the likelihood function.

b. Estimate parameters $\beta_0, \beta_1, \beta_2, \beta_3$ by maximum likelihood without using statistical software.

c. Find the standard errors for parameter estimates "by hand" using the Fisher information matrix and computational tools.

d. Test the hypothesis $H_0 : \beta_1 = 0$ versus $H_1 : \beta_1 \neq 0$. Compare your p-values with p-values given by statistical software and comment.

e. What are your findings about the influence of covariates on the response.