Exercise 1:

Consider the following binary model

$$P(Y_i = y_i) = \pi_i^{y_i}(1-\pi_i)^{1-y_i} \tag{1}$$

where

$$\pi_i = \frac{exp\{\mathbf{x}_i'\boldsymbol{\beta}\}}{1+exp\{\mathbf{x}_i'\boldsymbol{\beta}\}} \tag{2}$$

a) Simulate the data using the following values: - $n = 1000$ - $x_{0i} = 1 \,\forall\, i$, $x_{1i} \sim \mathcal{U}(18, 60)$, $x_{2i} \sim \mathcal{B}(0.5)$. - $\beta_0 = 1$, $\beta_1 = 0.1$, $\beta_2 = 1$.

b) Write down the likelihood function and the log-likelihood function in R as defined in the lecture.

c) Plot the likelihood function and the log-likelihood function for a range of values for the two parameters separately and show that they are maximized at the same value.

d) Estimate $\beta_0, \beta_1, \beta_2$ via maximum likelihood and calculate the standard errors. Use the estimation template provided in the lecture.

e) Propose and calculate a suitable method for the interpretation of the coefficients as discussed in the lecture.

f) Visualize your results, including the associated uncertainty around the estimates.

Exercise 2 (Simulation Study)**:**

a) Calculate the average training and the average prediction error for 100 simulation runs.

b) Consider the way the data is generated: propose and implement a change in the data generating process that (by your conjecture) will lead to an increase in the test error rate, but a decrease in the training error rate.

c) Bonus: Consider your assignment in 1 c): propose changes in the data generating process above that would make the likelihood function less informative over the range of parameter values. Implement these changes and show your results graphically.