Lecture Summary: March 13, 2023

- Logistic regression

  1. Simple logistic regression model

  Responses $(Y_i, x_i), i = 1, \ldots, n$. Y_i binary (i.e., $Y_i = 0$ or 1).

  Assume $\text{logit}(p_i) = \beta_0 + \beta_1 x_i,\ i = 1, \ldots, n$, where $p_i = \text{P}(Y_i = 1)$ and $\text{logit}(p) = \log[p/(1-p)]$ (natural logarithm).

  Solve for p_i: $p_i = e^{\beta_0+\beta_1 x_i}/(1 + e^{\beta_0+\beta_1 x_i})$.

  <u>Log-likelihood function</u>

  Likelihood: $L = \prod_{i=1}^n f(Y_i)$, where $f(Y_i) = p_i^{Y_i}(1-p_i)^{1-Y_i}$.

  Log-likelihood:

$$\begin{aligned}
\log L &= \sum_{i=1}^n \log[f(Y_i)] \\
&= \sum_{i=1}^n \{Y_i \text{logit}(p_i) + \log(1-p_i)\} \\
&= \sum_{i=1}^n \{Y_i(\beta_0 + \beta_1 x_i) - \log(1 + e^{\beta_0+\beta_1 x_i})\}.
\end{aligned}$$

  <u>Maximum likelihood</u>

  Find $\hat{\beta}_0, \hat{\beta}_1$ that maximize the log-likelihood over β_0, β_1.