Week 6 exercises

Michele De Vita

April 12, 2018

1. The linear probability model for binary response is simply $P(y = 1) = \pi_i = \mathbf{x_i'}\beta = \beta_{0i} + \beta_{1i}x_{1i} + \cdots + \beta_{ni}x_{ni}$ and require the restriction that $0 \leq \mathbf{X}'\beta \leq 1$

2. In the GLM we combine the probability output to the linear prediction through a function h called *response function* that it is a cumulative distribution function with co domain in [0,1]. In formula we can express the GLM as

$$P(y = 1) = \pi_i = h(\eta_i) = h(\mathbf{x_i'}\beta) = h(\beta_{0i} + \beta_{1i}x_{1i} + \cdots + \beta_{ni}x_{ni})$$

   $g = h^{-1}$ is the *link function* and it is used to calculate the linear predictor in function of probability: $\eta_i = g(\pi_i)$ The logit model use as response function the logistic function:

$$\pi = h(\eta) = \frac{e^{\eta}}{1 + e^{\eta}}$$

   . The linear predictor returns the log odds

$$\mathbf{x_i'}\beta = \beta_{0i} + \beta_{1i}x_{1i} + \cdots + \beta_{ni}x_{ni} = \pi_i = \log\left(\frac{\pi}{1 - \pi}\right)$$

   . The probit model use instead a normal distribution cumulative function.